Paula Silva Felicio

CARACTERIZAÇÃO MOLECULAR DE FAMÍLIAS DE ALTO RISCO PARA CÂNCER DE MAMA E/OU OVÁRIO HEREDITÁRIO, NEGATIVAS PARA VARIANTES GERMINATIVAS PATOGÊNICAS NOS GENES *BRCA1*/*BRCA2*

Tese apresentada ao Programa de Pós-Graduação da Fundação Pio XII - Hospital de Câncer de Barretos para obtenção do título de Doutor em Ciências da Saúde.

Área de concentração: Oncologia

Orientadora: Profa. Dra. Edenir Inêz Palmero

Coorientadores: Profa. Dra. Adriane Feijó Evangelista e
Prof. Dr. Matias Eliseo Melendez

Barretos, SP

2019

Paula Silva Felicio

CARACTERIZAÇÃO MOLECULAR DE FAMÍLIAS DE ALTO RISCO PARA CÂNCER DE MAMA E/OU OVÁRIO HEREDITÁRIO, NEGATIVAS PARA VARIANTES GERMINATIVAS PATOGÊNICAS NOS GENES *BRCA1*/*BRCA2*

Tese apresentada ao Programa de Pós-Graduação da Fundação Pio XII - Hospital de Câncer de Barretos para obtenção do título de Doutor em Ciências da Saúde.

Área de concentração: Oncologia
Orientadora: Profa. Dra. Edenir Inêz Palmero
Coorientadores: Profa. Dra. Adriane Feijó Evangelista e
Prof. Dr. Matias Eliseo Melendez

Barretos, SP
2019

FICHA CATALOGRÁFICA

F314c Felicio, Paula Silva.

Caracterização molecular de famílias de alto risco para câncer de mama e/ou ovário hereditário, negativas para variantes germinativas patogênicas nos genes BRCA1/BRCA2. / Paula Silva Felicio. - Barretos, SP 2019.

302 f. : il.

Orientador: Edenir Inêz Palmero.

Coorientadora: Adriane Feijó Evangelista / Matias Eliseo Melendez.

Tese (Doutorado em Ciências da Saúde) – Fundação Pio XII – Hospital de Câncer de Barretos, 2019.

1. BRCAX. 2. aCGH. 3. Exoma. 4. Câncer de mama e ovário hereditários. 5. Síndrome de predisposição hereditária ao câncer de mama e ovário. 6. BRCA1/BRCA2-negativas. I. Autor. II. Palmero, Edenir Inêz. III. Evangelista, Adriane Feijó / Melendez, Matias Eliseo. IV. Título.

CDD 616.994

FICHA CATALOGRÁFICA

Preparada por Martins Fideles dos Santos Neto CRB 8/9570
Biblioteca da Fundação Pio XII – Hospital de Câncer de Barretos

SUPORTE À PESQUISA POR AGÊNCIA DE FOMENTO

A tese de doutorado aqui apresentada foi desenvolvida no âmbito de financiamento pelo Hospital de Câncer de Barretos, Fundação Pio XII, por meio do Auxílio de Apoio ao Pesquisador (PAIP), pela Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP (processo número: 2013/24633-2), bolsa individual pelo Programa de Suporte à Pós-Graduação de Instituições de Ensino Particulares – PROSUP/CAPES e bolsa individual pelo Programa de Doutorado Sanduíche no Exterior – PDSE/CAPES (processo número: 88881.134079/2016-01).

As opiniões, hipóteses e conclusões ou recomendações expressas neste material são de responsabilidade dos autores e não necessariamente refletem a visão da FAPESP/CAPES.

DECLARAÇÃO DE RESPONSABILIDADE DOS PESQUISADORES

"Esta tese foi elaborada e está apresentada de acordo com as normas da Pós-Graduação do Hospital de Câncer de Barretos – Fundação Pio XII, baseando-se no Regimento do Programa de Pós-Graduação em Oncologia e no Manual de Apresentação de Dissertações e Teses do Hospital de Câncer de Barretos. Os pesquisadores declaram ainda que este trabalho foi realizado em concordância com o Código de Boas Práticas Científicas (FAPESP), não havendo nada em seu conteúdo que possa ser considerado como plágio, fabricação ou falsificação de dados. As opiniões, hipóteses e conclusões ou recomendações expressas neste material são de responsabilidade dos autores e não necessariamente refletem a visão da Fundação Pio XII – Hospital de Câncer de Barretos".

Embora o Núcleo de Apoio ao Pesquisador do Hospital de Câncer de Barretos tenha realizado as análises estatísticas e orientado sua interpretação, a descrição da metodologia estatística, a apresentação dos resultados e suas conclusões são de inteira responsabilidade dos pesquisadores envolvidos.

Os pesquisadores declaram não ter qualquer conflito de interesse relacionado a este estudo.

Dedico este trabalho aos meus pais Maria Cristina e Moacir (in memoriam), com todo o meu amor e gratidão. À minha avó Ruth, pelo imenso amor e pelas orações. À minha orientadora, Prof. Dra. Edenir, que tanto contribuiu e me deu diversas oportunidades para tornar este trabalho possível.

AGRADECIMENTOS

Foram muitas as pessoas que estiveram ao meu lado durante essa caminhada. Talvez eu não consiga expressar toda a minha gratidão por meio de palavras...

Antes de tudo, quero agradecer à **Deus** e aos **guias espirituais**, por ter abençoado todos os dias da minha vida, por iluminar meu caminho, me dar forças para seguir sempre em frente, por tantas oportunidades de aprendizado e por colocar na minha vida pessoas tão especiais.

À minha família, em especial aos meus pais, **Maria Cristina** e **Moacir** (*in memoriam*), e a minha **avó Ruth**, por todos os ensinamentos, investimentos e incentivos para que eu chegasse até aqui, por todo carinho, amor, dedicação e paciência.

Agradeço à minha orientadora, **Prof. Dra. Edenir Palmero**, pelos ensinamentos, por seu exemplo de cidadania e ética, por sua dedicação à profissão, por confiar na minha capacidade, pela amizade, carinho, cuidado e, sobretudo, sua paciência em todos os momentos. Não tenho palavras para expressar meus sentimentos! Tenho muito orgulho de dizer que você é minha orientadora! Você é um espelho como ser humano. Muito obrigada, por tudo!

Aos meus coorientadores, **Dr. Matias Melendez** e **Dra. Adriane Evangelista**, pelo auxílio e sugestões para o desenvolvimento deste trabalho.

I want to thanks **Dr. Patricia Tonin** and her research group at Research Institute – McGill University Health Centre for their support during my Ph.D. sandwich period in Montreal. I miss you guys (**Wejdan, Caitlin, Suzanna, and Supriya**)! I will never forget the great group you are!

Agradeço as oportunidades que tive no Canadá no campo pessoal e profissional. Ainda, agradeço as amigas **Iza** e **Marcela** que estiveram comigo durante esse período em Montreal, sem vocês essa experiência não teria sido tão especial!

Aos colaboradores do Centro Internacional de Pesquisa (CIPE) do Hospital A.C. Camargo Cancer Center: **Dra. Dirce Maria Carraro**, **Dra. Bruna Durães de Figueiredo Barros** e **Dra. Giovana Tardin Torrezan** pelo auxílio envolvendo a análise do sequenciamento exômico.

Ao **Dr. Luís Felipe Pinto** e à **Dra. Sheila Coelho Soares Lima**, do Instituto Nacional do Câncer, por me receberem de forma tão agradável e por todo conhecimento que foi me passado.

Aos membros da Banca de Acompanhamento, **Dra. Patricia Asthon-Prolla** e **Dr. Rui Manuel Vieira Reis**, pelas considerações e pontos positivos adicionados a este trabalho.

À toda a equipe do **Departamento de Oncogenética** pelas informações fornecidas, em especial ao **Dr. Henrique Galvão**.

Ao **Centro de Diagnóstico Molecular**, em especial à **Gabriela**, **Cristina**, **André**, **Sâmia** e **Gustavo**, pela ajuda, ensinamentos e também pelos resultados moleculares das pacientes incluídas no estudo.

À todos os alunos do **Grupo de Pesquisa em Câncer Hereditário** do Hospital de Câncer de Barretos, pelo companheirismo, apoio e incentivo, em especial à **Natalia** e **Rebeca** pela ajuda em revisar os prontuários clínicos e os dados de história familiar.

Aos **pacientes** e **seus familiares**, que permitiram todo e qualquer conhecimento gerado.

Ao **Departamento de Patologia** do Hospital de Câncer de Barretos, em especial ao **Dr. Cristovam Scapulatempo-Neto**, à **Dra. Lígia Maria Kerr** e ao **Dr. Marcus de Medeiros Matsushita** pela avaliação histopatológica das amostras. À **Bia** e **Paty** do arquivo de material, pela imensa dedicação na busca dos casos e corte dos blocos.

Aos **biologistas do CPOM** pela disponibilidade em nos ajudar.

Ao **Programa de Pós-Graduação** do Instituto de Ensino e Pesquisa do Hospital de Câncer de Barretos, o qual me proporcionou um enorme aprendizado.

À **Brenda**, **Silvana**, **Cíntia**, **Mariana** e **Dely** da Pós-Graduação, pela dedicação e competência em ajudar os alunos.

Aos bibliotecários, **Martins** e **Milene**, pelo modo prestativo e eficiente com que sempre atenderam às minhas necessidades.

Às amigas **Ana Laura**, **Aline Rocha**, **Estela**, **Fer Cury**, **Natalia Campacci** e **Rebeca** pela amizade, apoio, carinho e conselhos nas horas difíceis. Fico feliz com a determinação e o entusiasmo de vocês. E torço para que alcancem seus objetivos. Obrigado pelos momentos de alegria que passamos juntas e pelo convívio maravilhoso! Pelo carinho e ajuda que sempre recebi.

À **Carolzinha** e **Letícia** por todas as palavras de carinho e suporte. Vocês são um exemplo de determinação, força, generosidade e bondade!

A todos os **amigos do CPOM** pela convivência e troca de experiências. Obrigada especial para àqueles que sempre me apoiaram e incentivaram: **Weder**, **Yollanda**, **Murilo**, **Maraísa**, **Rhafa**, **Ariane**, **Allini Mafra**.

Às agências de fomento à pesquisa: **CAPES** e **FAPESP**.

Agradeço à **Banca Julgadora** pelo tempo atribuído à leitura desta tese.

Às demais pessoas que de uma maneira ou de outra contribuíram na elaboração deste trabalho ou participaram da minha vida, e que, por ventura, eu tenha me esquecido de agradecer. O meu sincero agradecimento!

“É exatamente disso que a vida é feita: de momentos! Momentos os quais temos que passar, sendo bons ou não, para o nosso próprio aprendizado, por algum motivo. Nunca esquecendo do mais importante: nada na vida é por acaso.”

Chico Xavier

ESBOÇO DA TESE

A tese de doutorado aqui apresentada foi elaborada no modelo baseado em trabalho publicado. Para essa modalidade de apresentação, as seções da “Introdução” e “Discussão” são sucintas e tem o objetivo de situar e atualizar o leitor sobre o assunto. Além disso, as seções “Material e Métodos” e “Resultados” foram substituídas pelos dos artigos publicados ou a serem publicados.

SUMÁRIO

LISTA DE FIGURAS

Estudo de alterações no número de cópias cromossômicas em amostras tumorais de mulheres com história pessoal e familiar sugestiva de câncer de mama hereditário, através de aCG.

Genetic alterations detected by comparative genomic hybridization in BRCAX breast and ovarian cancers of Brazilian population.

Estudo de variantes germinativas patogênicas nas amostras BRCA1/BRCA2-negativas, através do sequenciamento exômico

Whole-exome sequencing of Brazilian non-*BRCA1*/*BRCA2* mutation carrier cases at high-risk for hereditary breast/ovarian cancer.

Candidate DNA repair susceptibility genes identified by whole-exome sequencing in non-*BRCA1*/*BRCA2* carriers at high risk for hereditary breast/ovarian cancer.

LISTA DE TABELAS

Estudo da presença de variante germinativa fundadora portuguesa no gene *BRCA2* (c.156_157insAlu), através da PCR seguida do sequenciamento de Sanger

Estudo de alterações no número de cópias cromossômicas em amostras tumorais de mulheres com história pessoal e familiar sugestiva de câncer de mama hereditário, através de aCG

Estudo de variantes germinativas patogênicas nas amostras BRCA1/BRCA2-negativas, através do sequenciamento exômico

LISTA DE ABREVIATURAS

aCGH	*Array*-CGH
CM	Câncer de mama
CO	Câncer de ovário
HBOC	*Hereditary Breast and Ovarian Cancer Syndrome*
LOH	*Loss of Heterozygosity*
WES	*Whole-Exome Sequencing*
WGS	*Whole-Genome Sequencing*

RESUMO

Estima-se que aproximadamente 10% de todos os tumores sejam de caráter hereditário e, ainda, 20-30% correspondam a agrupamentos de tumores ocorridos em uma mesma família, ou tumores familiais. A identificação de indivíduos em risco para câncer hereditário é importante, pois: (1) indivíduos afetados apresentam risco cumulativo vital elevado para o desenvolvimento de câncer; (2) familiares de um indivíduo afetado podem estar em risco; e (3) existe medidas de rastreamento intensivo e intervenção preventiva que podem diminuir, significativamente, o risco de câncer em portadores de variantes patogênicas. No entanto, apesar de existirem *loci* gênicos associados a uma maior predisposição/suscetibilidade à Síndrome de Predisposição Hereditária ao Câncer de Mama e Ovário (HBOC), para uma grande parcela dos casos (cerca de 50%), a causa genética permanece uma incógnita. Portanto, o objetivo central do estudo foi realizar uma caracterização molecular de um grupo de 52 mulheres não aparentadas (52 famílias) com alto risco para HBOC, negativas para variantes patogênicas nos genes *BRCA1* e *BRCA2*. Foram empregadas estratégias de (1) análise pontual de variantes não identificadas pelas técnicas de rotina empregadas nas análises de *BRCA1*/*BRCA2* (detecção da inserção ALU no gene *BRCA2*), através de PCR seguido de eletroforese, (2) análise da presença de alterações cromossômicas (DNA tumoral) por meio da plataforma *Agilent Technologies - SurePrint G3 Unrestricted CGH 8x60K* e (3) sequenciamento exômico de DNA germinativo, através da plataforma *NextSeq 500* (*Illumina*). Dessa forma, foi identificada a variante c.156_157insAlu em *BRCA2* em 0,65% dos pacientes avaliados (9/1.380), sendo que, nesses portadores foi observada uma maior proporção da ancestralidade Européia (80%), seguida da Africana (10%) e Ameríndia. Além disso, na maioria das famílias com a inserção Alu identificadas, o haplótipo detectado é compatível com o haplótipo ancestral português. No que se refere à análise dos dados de rearranjos cromossômicos, foram identificadas 20 regiões com ganhos genômicos e 31 com perdas no tecido tumoral das pacientes analisadas. Foi observada uma associação da perda de 22q13.31-13.32 e a presença de casos de câncer de ovário. Entre os genes presentes nas regiões comumente alteradas, encontramos *FGFR1*, e *MIR3201*. Em relação à análise de alterações gênicas no tecido germinativo, por meio do sequenciamento exômico, foi observado um total de 53 variantes com perda de função e 128 variantes do tipo *missense*. Destes genes, 23 foram

descritos como "*hallmarks of cancer*" pelo banco de dados COSMIC, destacando os genes *ATM*, *CHEK2*, *PMS2*, *KRAS*, *KIT* e *EGFR*. Além disso, foram observadas variantes envolvendo a família RAD, com destaque para os genes *RAD50*, *RAD51C* e *RAD54L*. Quando analisamos as vias de reparo mais comumente alteradas, observamos um enriquecimento nas vias de reparo de Recombinação Homóloga, *Mismatch Repair* e Anemia de Fanconi. Em conclusão, este é o maior estudo brasileiro envolvendo famílias de alto risco para HBOC, que utilizou diversas técnicas de larga escala para a caracterização dessas famílias e identificação do fator causal/associado ao "câncer de mama hereditário do tipo x". Estes dados sugerem novos *insights* sobre os fatores genéticos envolvidos no câncer de mama e ovário hereditários e reforçam a necessidade de ampliarmos nossa estratégia de teste molecular para as famílias de alto risco e negativas para variantes patogênicas nos genes *BRCA1/BRCA2*.

PALAVRAS-CHAVE: BRCAX, aCGH, exoma, câncer de mama e ovário hereditários, síndrome de predisposição hereditária ao câncer de mama e ovário.

1. INTRODUÇÃO

1.1. Câncer – Aspectos Gerais

O câncer é considerado um problema de saúde pública tanto para os países desenvolvidos quanto para as nações em desenvolvimento. O câncer é responsável por mais de 12% de todas as causas de óbito no mundo: mais de sete milhões de pessoas morrem anualmente da doença[1]. Como a expectativa de vida no planeta tem melhorado gradativamente, a incidência de câncer, estimada em cerca de 12 milhões de casos em 2008, deverá superar a marca de 20 milhões em 2025[2]. No Brasil as estimativas para o ano de 2018 (válidas também para o ano de 2019) apontam para a ocorrência de mais de 600 mil novos casos de câncer[3].

O câncer é uma doença multifatorial resultante do acúmulo de alterações genéticas e epigenéticas que levam ao crescimento desordenado de células anormais[4], à alterações no ciclo celular e ao acúmulo de erros no DNA[5,6]. Tais modificações foram descritas por Hanahan e Weinberg[7], como "*Hallmarks of Cancer*" e incluem importantes processos de controle do ciclo celular, do metabolismo energético e do sistema imune (Figura 1). Quando estes mecanismos são afetados e não ocorre o reparo, células anormais geram descendentes que herdam a propensão para proliferar sem responder à regulação, o que resulta em uma proliferação celular clonal capaz de se expandir indefinidamente. Esse processo de proliferação celular descontrolada pode gerar uma produção demasiada de células, podendo originar o câncer[8,9].

Figura 1 – Propriedades básicas adquiridas pelas células tumorais. (Fonte: adaptado de Hanahan e Weinberg[7]).

Em relação aos genes envolvidos no processo de carcinogênese existem duas classes principais: os oncogenes e os genes supressores tumorais. Os oncogenes são formas alteradas dos proto-oncogenes, genes cujas funções principais estão relacionadas à regulação do crescimento, proliferação e diferenciação de células normais. Para a ativação dos oncogenes basta que um dos dois alelos esteja alterado, o que pode desencadear o processo de desenvolvimento tumoral[10]. A ativação dos oncogenes pode ocorrer através de diferentes processos como: amplificação gênica, variante pontual e translocação cromossômica[10,11]. Por outro lado, os genes supressores tumorais são genes que codificam proteínas envolvidas na manutenção da estabilidade genômica e na proliferação celular e, portanto, a perda da função destes genes pode permitir o desenvolvimento tumoral[6]. Segundo o modelo proposto por Knudson, para que ocorra a transformação neoplásica é necessário que haja a inativação dos dois alelos de um gene supressor de tumor, ao contrário do que ocorre com os oncogenes[12]. Os principais mecanismos responsáveis pela inativação dos genes supressores tumorais incluem: perda de heterozigosidade (LOH – *loss of heterozygosity*), presença de variantes patogênicas e alterações no padrão de metilação da região promotora[13,14].

1.2. Câncer de Mama

De acordo com o GLOBOCAN, o câncer de mama é a neoplasia maligna mais frequente no sexo feminino, representando 24% dos casos novos de câncer a cada ano[1]. No Brasil, o câncer de mama se constitui na principal causa de incidência e mortalidade por câncer entre as mulheres. Conforme o Instituto Nacional de Câncer (INCA) estimam-se 57.700 novos casos de câncer de mama, para cada ano do biênio 2018-2019[3].

Estima-se que para o câncer de mama, assim como para grande parte dos tumores malignos conhecidos, 10-15% sejam de caráter hereditário[15,16], e 20-30% correspondam a agrupamentos de tumores ocorridos em uma mesma família, ou tumores familiais[17].

1.3. Câncer de Ovário

Mundialmente, o câncer de ovário (CO) é considerado o oitavo tumor mais frequente no sexo feminino. Segundo o GLOBOCAN, em 2018, aproximadamente 295.414 novos casos de CO seriam relatados, representando cerca de 6% dos novos casos de câncer nas mulheres, sendo o oitavo tumor mais incidente no mundo[1]. A estimativa do INCA para o biênio 2018-2019 é de 6.150 novos casos por ano, ocupando o oitavo lugar com maior incidência entre as mulheres no Brasil[3].

Estima-se para o CO, que cerca de 23% dos casos têm relação com a hereditariedade, ou ainda, apresentam agrupamentos familiares[18,19].

1.4. Câncer de Mama e Ovário Hereditário

A primeira descrição de uma família com câncer de mama hereditário foi publicada em 1866, pelo cirurgião francês Paul Broca. Broca descreveu detalhadamente quatro gerações de mulheres acometidas por câncer de mama na família de sua esposa, em que 10 de 24 mulheres foram afetadas pela doença. A história familiar relatada indicava pela primeira vez uma predisposição hereditária ao câncer de mama[20]. Em 1976, David Anderson relatou que mulheres com história de vários familiares de primeiro grau com câncer de mama possuíam Risco Cumulativo Vital (RCV) de desenvolver a doença 47 a 51 vezes maior que o risco da população em geral. Ainda, referiu que nessas mulheres o tumor se desenvolvia, em geral, antes da menopausa, era bilateral e estava associado aos ovários[21]. Diversos estudos foram

realizados e confirmaram as observações iniciais: a existência de uma predisposição aumentada ao câncer de mama de acordo com a história familiar[22–24].

Com o avanço do conhecimento e das técnicas de biologia molecular foi possível identificar genes que, quando alterados, aumentam significativamente o risco de desenvolver diversos tumores, incluindo o câncer de mama e ovário, se destacando os genes supressores tumorais *BRCA1* e *BRCA2*[25–27]. Acredita-se que os genes *BRCA1* e *BRCA2* sejam responsáveis por aproximadamente 25% de todos os casos de câncer de mama e ovário hereditário[17,28–30]. Outros genes tais como *TP53*, *PTEN*, *STK11* e *CDH1* estão associados a um alto risco de desenvolvimento de câncer de mama, provavelmente devido ao fato de que, nas síndromes em que estão envolvidos, o câncer de mama faz parte do espectro tumoral[29]. Além disso, existe um grupo de genes associado a um aumento moderado no risco para o câncer de mama, estando esses genes envolvidos principalmente em vias de reparo de danos no DNA, como *ATM*, *CHEK2*, *BRIP1* e *RAD51C*. E, por último, existem genes de baixo risco, cuja frequência do menor alelo pode ser de até 5%, e que coletivamente ocasionam um pequeno aumento na susceptibilidade ao câncer de mama[31,32].

No entanto, apesar dos avanços nas técnicas de biologia molecular, os genes até agora associados à Síndrome de Predisposição Hereditária ao Câncer de Mama e Ovário (do inglês: *Hereditary Breast and Ovarian Cancer Syndrome* - HBOC), sejam eles de alto, moderado ou baixo risco, respondem por aproximadamente 50% dos casos[17,29]. Dessa forma, para a uma significativa parcela de famílias em risco para câncer de mama e/ou ovário hereditários, nenhuma variante germinativa patogênica é detectada e, dessa maneira, informações acerca do risco de desenvolvimento de câncer para outros membros da família, medidas preventivas e redutoras de risco acabam por se basear em uma quantidade limitada de informações e tendem a ser genéricas[33].

1.4.1. Genes de predisposição ao câncer de mama e/ou ovário

Os genes de predisposição hereditária ao câncer de mama e ovário podem ser categorizados de acordo com seu risco relativo, ou seja, a chance de uma determinada alteração em um gene aumentar o risco de desenvolvimento de um determinado tipo de câncer. De acordo com a risco, os genes podem ser categorizados em: genes de alto risco (RR: maior do que 5), moderado (RR: entre 1,5 a 5) e baixo (RR: menor que 1,5)[34].

1.4.1.1. Genes de alto risco

1.4.1.1.1. *BRCA1*

O gene *BRCA1* (do inglês: *breast cancer 1*) localizado no cromossomo 17 (17q21.3), codifica uma proteína multifuncional com 1.863 aminoácidos, apresentando um papel fundamental na manutenção da estabilidade genómica. Além disso, está envolvido em muitos processos celulares importantes na biologia tumoral incluindo o reparo do DNA, a progressão do ciclo celular e a regulação da transcrição[35]. Com exceção de algumas populações com efeito fundador conhecido associado ao gene *BRCA1*, tal como os judeus Ashkenazi, variantes patogênicas associadas a "*hotspots*" em *BRCA1* não são usualmente encontradas, porém, dado o fato do éxon 11 apresentar 3.426 pb (mais de 50% da região codificante do gene), a maioria das variantes são encontradas neste éxon[36]. Estudo realizado pelo nosso grupo de pesquisa envolvendo 349 mulheres brasileiras não aparentadas entre si, identificou que uma alteração no éxon 11 (c.3331_3334delCAAG) e uma no éxon 20 (c.5266dupC) do gene *BRCA1* foram as alterações mais frequentes, representando aproximadamente 35% das variantes identificadas e mais de 50% das variantes patogênicas em *BRCA1*[37]. Além disso, outro estudo realizado por Palmero e colaboradores (2018) e que envolveu uma compilação de variantes patogênicas identificadas na população Brasileira, identificou 126 variantes patogênicas diferentes em *BRCA1*, presentes em 441 probandos. Ainda, dentre as variantes patogênicas mais prevalentes, a variante fundadora Europeia c.5266dupC foi a mais comum dentre os pacientes, correspondendo a 20,2% de todas as variantes encontradas no *BRCA1*[38]. Através dos dados do "*Consortium of Investigators of Modifiers of BRCA1/2*", Rebbeck *et al.* (2018) objetivou investigar a ocorrência de variantes patogênicas por família, sendo que um total de 18.435 famílias com variantes patogênicas em *BRCA1* foram analisadas. Por meio dessa análise, os autores identificaram 1.650 variantes únicas. Dentre as variantes mais frequentemente identificadas, os autores destacaram que as variantes fundadoras Europeia c.5266dup (5382insC) e c.68_69del (185delAG) foram observadas em todos os continentes. Ainda, a variante c.3331_3334del foi relatada como mais comumente identificada na América Central e do Sul[39].

Mulheres portadoras de variantes germinativas patogênicas em *BRCA1* apresentam alto risco de desenvolver câncer de mama e ovário. Os riscos ao longo da vida são de até 87% e 68%, respectivamente[40]. Em relação à histopatologia, os tumores relacionados ao *BRCA1* são geralmente caracterizados pela falta de expressão do receptor de estrógeno, receptor de

progesterona e HER2 (do inglês: *Human Epidermal growth factor Receptor-type 2*)[41]. Além do câncer de mama e ovário, portadores de variantes patogênicas no gene *BRCA1* também apresentam risco aumentado para o desenvolvimento de câncer de mama contralateral, câncer de próstata e colorretal[16,40].

Segundo as diretrizes do *National Comprehensive Cancer Network* (NCCN – versão 3.2019), diversas medidas são oferecidas para pacientes portadores de variantes patogênicas em *BRCA1*, dependendo da idade. O rastreamento com ressonância magnética e mamografia anual são medidas recomendadas. Além disso, as opções de mastectomia e salpingo-ooforectomia para redução de risco devem ser discutidas durante o aconselhamento genético[42].

1.4.1.1.2. *BRCA2*

O gene *BRCA2* (do inglês: *breast cancer 2*) codifica uma proteína que desempenha um papel crítico nas vias de reparo do DNA. Além disso, BRCA2 é responsável por recrutar a proteína RAD51 para mediar o reparo de quebras dupla fita do DNA[35].

Homens portadores de variantes germinativas patogênicas em *BRCA2* apresentam um risco ao longo da vida de desenvolver câncer de próstata, mama e pâncreas, em aproximadamente 20%, 7% e 6%, respectivamente[43–45]. Em relação ao risco de câncer para mulheres portadoras de variantes patogênicas em *BRCA2*, estudos apontam para um risco cumulativo vital de 40-65% para câncer de mama e 20% para câncer de ovário[40,46]. Os tumores relacionados a *BRCA2* normalmente são caracterizados pela expressão de receptores de estrógeno e progesterona, sendo comumente classificados como tumores do subtipo Luminal B, e ainda, tendem a ter características semelhantes aos tumores mamários esporádicos, ao contrário dos tumores relacionados à *BRCA1*[47,48]. Assim como o gene *BRCA1* que não apresenta "*hotspots*", o gene *BRCA2* também não apresenta sítios preferenciais de mutações ou "*hotspots*". No entanto, algumas variantes apresentam frequências aumentadas. Estudo realizado por Fernandes e colaboradores (2016), indicam que a variante germinativa c.2808_2811delACAA, localizada no éxon 11 do gene *BRCA2*, apresenta uma frequência de 13,6% dentre os probandos provenientes da população brasileira analisadas[37]. Adicionalmente, estudo publicado por Palmero *et al*. (2018), através da análise de 649 probandos brasileiros, identificou a presença de 208 variantes em *BRCA2,* sendo que, 103 variantes foram observadas apenas uma vez. Além disso, 17 variantes não descritas em

bancos de dados (ClinVar, BRCA Share, LOVD, ARUP e BRCA Exchange) foram identificadas. Ainda, os autores reportaram que a variante fundadora portuguesa (c.156_157insAlu) foi o rearranjo genômico mais comumente identificado na população de estudo[38]. No que refere ao cenário mundial, Rebbeck *et al.* (2018) analisou 11.351 famílias com variantes patogênicas em *BRCA2*, sendo que 1.731 variantes foram consideradas únicas, ou seja, observadas apenas em uma família. Ainda, os autores relataram uma alta frequência da variante fundadora Europeia (c.5946del) em famílias de todos os continentes. Adicionalmente, outras variantes como c.2808_2811del, c.2T>G e c.156_157insAlu também foram comumente observadas em famílias da América Central e do Sul[39].

Assim como mencionado anteriormente, dentre as medidas para o manejo dos pacientes com variantes patogênicas em *BRCA2*, segundo as diretrizes do NCCN, destacam-se a ressonância magnética e mamografia anual. As opções de mastectomia e salpingo-ooforectomia também devem ser discutidas para redução de risco, e ainda, a triagem de câncer de próstata para homens portadores de *BRCA2* deve ser oferecida.

1.4.1.1.3. *CDH1*

O gene *CDH1* (do inglês: *cadherin 1*) codifica a proteína E-caderina, um membro da família das glicoproteínas transmembrana. A E-caderina é expressa nos tecidos epiteliais e é responsável pela adesão célula-célula dependente do cálcio. A perda da expressão de *CDH1* foi associada ao aumento do potencial de invasão das células cancerígenas[35,49]. Variantes germinativas patogênicas em *CDH1* são associadas ao carcinoma gástrico difuso hereditário (do inglês: *Hereditary Diffuse Gastric Cancer – HDGC*)[50]. Aproximadamente 30% das famílias com HDGC apresentam na história familiar indivíduos com câncer de mama do tipo lobular[51]. Estima-se que os indivíduos heterozigotos para variantes patogênicas no gene *CDH1* tenham um risco cumulativo de câncer gástrico difuso superior a 80%, tanto em homens quanto em mulheres, até os 80 anos. Além disso, mulheres heterozigotas apresentam um risco cumulativo de câncer de mama lobular de 60% até os 80 anos de idade[51,52].

Variantes germinativas no gene *CDH1*, embora raras, foram descritas em mulheres diagnosticadas com câncer de mama lobular (CML) e sem história familiar de câncer gástrico difuso. Além disso, há evidências que o CML poderia ser a primeira manifestação de HDGC[53–55]. Estudo realizado por Corso e colaboradores (2016), analisou a presença de variantes germinativas no gene *CDH1* em 482 casos de CML. A história familiar de câncer de mama foi

relatada em 40,7% dos pacientes analisados e, ainda, 20,3% apresentaram câncer de mama bilateral[56].

Devido ao risco considerável de câncer de mama lobular em portadoras de variantes patogênicas em *CDH1*, as diretrizes do NCCN recomendam a triagem com mamografia anual, a partir dos 30 anos de idade. No entanto, o rastreamento pode ser iniciado mais cedo em pacientes com histórico familiar de câncer de mama de início precoce. Além disso, a mastectomia pode ser discutida, dependendo da história familiar de câncer[42].

1.4.1.1.4. *PALB2*

O gene *PALB2* (do inglês: *partner and localizer of BRCA2*) também conhecido como *FANCN*, é um gene associado à Anemia de Fanconi e codifica uma proteína que interage com BRCA2 durante a recombinação homóloga e o reparo de quebras dupla fita no DNA[35]. Variantes patogênicas em *PALB2* foram detectadas em 1-5% dos pacientes com câncer de mama hereditário e negativos para variantes patogênicas em *BRCA1/BRCA2*[57–59]. Além disso, alguns autores indicam que variantes patogênicas em *PALB2* estão associadas a um alto risco para o desenvolvimento de câncer de mama triplo-negativo (RR: 14,41 -IC 95%: 9,27 a 22,60)[60,61]. Em relação ao risco relativo para câncer de ovário entre portadores de variantes patogênicas em *PALB2*, os resultados ainda são controversos. Estudos descrevem que esse gene apresenta um risco baixo [1,60 (IC 95%: 0,98 a 2,60)[62] a moderado [2,31 (IC 95%: 0,77 a 6,97)][63] no desenvolvimento de câncer do ovário. Estudo realizado em 2011 por Casadei *et al.,* realizou o sequenciamento do gene *PALB2* em famílias de alto risco para HBOC e com pacientes diagnosticados com câncer de mama. Foram identificadas variantes germinativas patogênicas em 33 de 972 famílias (3,4%). Além disso, 18/33 famílias (55%) apresentavam pelo menos um familiar com câncer de ovário. Ainda, tais famílias tinham um espectro tumoral semelhante àquele das famílias com variantes germinativas patogênicas em *BRCA2*[64].

Em outro estudo, publicado em 2012 por Tischkowitz e colaboradores, variantes patogênicas em *PALB2* foram identificadas em mulheres com câncer de mama. Foi observado que parentes de primeiro grau e do sexo feminino das pacientes apresentaram um risco relativo mais elevado de desenvolver câncer de mama, quando comparadas àquelas em que não apresentavam variantes patogênicas na família, indicando conforme estimativas realizadas pelos autores do trabalho, que as variantes germinativas patogênicas em *PALB2* conferem um aumento de aproximadamente 5,3 vezes no risco[65].

Com base nos estudos citados anteriormente, através das diretrizes do NCCN, diversas recomendações para os portadores de variantes germinativas patogênicas em *PALB2* foram adotadas[42]. Têm sido recomendado que pacientes com variantes germinativas patogênicas nesse gene realizem mamografia anual, a partir dos 30 anos, e, além disso, a mastectomia profilática pode ser considerada[42].

1.4.1.1.5. *PTEN*

Variantes germinativas patogênicas no gene *PTEN* (do inglês: *phosphatase and tensin homolog*) são a causa da Síndrome de Cowden, caracterizada por múltiplos hamartomas, alto risco de desenvolver tumores benignos e malignos na tireóide, mama e endométrio. Também podem ser observadas lesões mucocutâneas, anormalidades da tireóide, doença fibrocística, leiomioma uterino múltiplo e macrocefalia[66,67]. Os indivíduos afetados têm um risco cumulativo vital de até 50% para câncer de mama, 10% para câncer de tireóide e 5-10% para câncer de endométrio[68,69]. Além disso, mais de 90% dos indivíduos com a Síndrome de Cowden expressarão alguma manifestação clínica até a segunda década de vida[70].

Segundo as diretrizes do NCCN, diversas medidas são recomendadas para a prevenção e a detecção precoce de tumores frequentemente associados à síndrome de Cowden, incluindo exames físicos anuais, a partir dos 18 anos de idade (ou ainda, 5 anos antes da idade mais jovem do diagnóstico de um familiar). As recomendações para mulheres portadoras de variantes patogênicas em *PTEN* incluem a realização anual de mamografia e ressonância magnética e, ainda, a mastectomia profilática deve ser oferecida. Ainda, vale destacar que a ooforectomia não é indicada, porém é necessário avaliar outros motivos que levem à necessidade de realizar essa cirurgia[42].

1.4.1.1.6. *STK11*

O gene *STK11* (do inglês: *serine/threonine kinase 11*), ou também conhecido como *LKB1*, codifica uma serina-treonina quinase que está envolvida na regulação do metabolismo, diferenciação celular, proliferação e apoptose[35]. Variantes no gene *STK11* foram associadas à Síndrome de Peutz-Jehgers (SPJ), a qual é caracterizada por pólipos hamartomatosos no trato gastrointestinal, lesões mucocutâneas pigmentadas e predisposição ao câncer[71]. Dentre os tumores associados à SPJ podemos destacar: colorretal, gástrico, mama, ovário, útero, pâncreas, pulmão, testículos, bem como pâncreas[72,73]. Portadores de variantes patogênicas

no gene *STK11* apresentam um risco aumentado de desenvolver qualquer câncer mencionado acima, sendo que, o RCV para o desenvolvimento de cada tipo tumoral é de 39% para o câncer colorretal, 29% para o câncer gástrico, 24-54% para câncer de mama, 21% para o ovário, 10-23% para o câncer do útero, 9% para o câncer testicular, 7-17% para o câncer de pulmão e 11-36% para o câncer de pâncreas[73,74].

Dentre as medidas que podem ser oferecidas para pacientes portadores de variantes patogênicas em *STK11*, segundo as diretrizes do NCCN, estão a mamografia e ressonância magnética, enterografia do intestino delgado por tomografia computadorizada ou ressonância magnética, além do exame pélvico, Papanicolau e ultrasson transvaginal. Devido a presença de poucos dados sobre o benefício de mastectomia profilática para a redução de risco em mulheres com variantes patogênicas em *STK11*, as diretrizes do NCCN não recomendam essa prática nos pacientes, porém esse procedimento pode ser considerado com base na história familiar de câncer[42].

1.4.1.1.7. *TP53*

O gene *TP53* (do inglês: *tumor protein p53*) codifica um fator de transcrição que está envolvido nas respostas celulares ao estresse ambiental e genotóxico[35].

A Síndrome de Li-Fraumeni (SLF) é causada por variantes germinativas patogênicas no gene *TP53*, sendo caracterizada por predisposição a múltiplos tumores de início precoce. A SLF apresenta alta variabilidade na penetrância, idade de diagnóstico e espectro tumoral[75]. Os portadores de variante germinativa em *TP53* apresentam um RCV superior a 90% para o desenvolvimento de câncer[76–78]. Os tumores mais comuns associados à SLF são sarcoma , câncer de mama , tumores cerebrais e câncer adrenocortical . Outros tumores incluem colorretal, pâncreas, leucemia e gástrico[79,80]. Embora a SLF represente uma pequena fração de casos de câncer de mama (~0,1%), os portadores dessa síndrome têm 18 a 60 vezes mais risco de desenvolver esse tumor com idade precoce (antes dos 45 anos) quando comparados com a população em geral[81–83].

As recomendações para o rastreamento do câncer de mama em portadores de variantes patogênicas em *TP53* incluem ressonância magnética ou mamografia. Além disso, a realização de mastectomia profilática pode ser considerada, bem como, outros exames podem ser solicitados de acordo com a história pessoal e família de câncer, tais como ressonância magnética de todo o corpo, colonoscopia e endoscopia[42].

1.4.1.2. Genes de moderado risco

1.4.1.2.1. *ATM*

A proteína produzida pelo gene *ATM* (do inglês: ATM *serine/threonine kinase*) é uma proteína quinase relacionada à via PI3K, apresentando várias funções complexas, incluindo papel central no reparo de quebras dupla fita no DNA, juntamente com as proteínas codificadas pelos genes *TP53*, *BRCA1* e *CHEK2*[35].

Variantes germinativas bialélicas em *ATM* causam Ataxia Telangiectasia (A-T), que confere um alto risco para várias patologias hematológicas, principalmente em idade precoce. Além disso, portadores de variantes germinativas em heterozigose são encontrados em cerca de 0,35 – 1% da população geral, e apresentam um risco aumentado para o desenvolvimento de câncer de mama (2,4 vezes)[72,84]. Um estudo publicado em 2011 por Goldgar e colaboradores, demonstrou um risco significativamente aumentado de câncer de mama em portadores de variantes que levam à perda de função proteica com uma penetrância semelhante à conferida por variantes germinativas em *BRCA2*[85]. Além disso, através do estudo publicado por Couch *et al.* (2017), foi observado que mulheres portadoras de variantes germinativas em heterozigose em *ATM* apresentaram um maior risco de câncer de mama, com um RR estimado de 2,78 (IC 95%: 2,22-3,62)[86]. Em relação ao risco de desenvolvimento de câncer de ovário, Kurian *el al* (2017), relataram um baixo risco associado à presença de variantes patogênicas em *ATM* (1,92 [IC 95%: 0,92 a 4,23])[62].

Para as portadoras de variantes patogênicas em *ATM*, o NCCN indica a realização anual de mamografia ou ressonância magnética. A realização de matectomia profilática pode ser considerada com base na história familiar de câncer. Em elação à triagem do câncer de ovário, atualmente não há evidências suficientes para a recomendação da salpingo-ooforectomia redutora de risco[42].

1.4.2.1.2. *BRIP1*

O gene *BRIP1* (do inglês: *BRCA1 interacting protein C-terminal helicase 1*) codifica uma proteína que interage com BRCA1[35].

Em 2006, variantes que levam à produção de uma proteína truncada em *BRIP1* foram identificadas em famílias com câncer de mama e sem variantes germinativas conhecidas em

outros genes[87]. Embora haja relatos de maiores riscos em algumas famílias, o risco relativo de câncer de mama foi estimado em torno de 2 vezes[62,86]. Além disso, estudos realizados por Couch *et al*. (2015), Buys *et al*. (2017) e Shimelis (2018), observaram que a prevalência de variantes patogênicas em *BRIP1* foi maior entre as mulheres com câncer de mama triplo-negativo (RR: 2,46 (IC 95%: 1,18 to 5,08)[60,88,89].

De maneira adicional, foi descrito que variantes germinativas em *BRIP1* também conferem um risco aumentado de câncer de ovário[62,90]. Estudo publicado por Ramus *et al.* (2015), avaliando 3.236 mulheres com câncer de ovário epitelial, relatou um risco relativo de 11,2 (IC 95%: 3,22-34,1) para o desenvolvimento de câncer de ovário[91]. Ainda, estudo conduzido por Weber-Lassalle e colaboradores (2018), observou que variantes patogênicas em *BRIP1* conferem um alto risco (OR: 20,9; IC 95%: 12,0-36,5) para o desenvolvimento de câncer de ovário, especialmente com início tardio (média ao diagnóstico de 61 anos, variando de 26 a 76 anos; OR: 29,9; IC 95%: 14,9-59,6)[92].

Em relação às diretrizes do NCCN, recomenda-se que a realização da salpingo-ooforectomia redutora de risco em portadores de variantes patogênicas em *BRIP1* seja considerada a partir dos 45 anos de idade, de acordo com a história pessoal e familia de câncer. Além disso, é recomendado que as mulheres realizem mamografia a partir dos 40 anos de idade[42].

1.4.1.2.3. *BARD1*

O gene *BARD1* (do inglês: *BRCA1 associated RING domain 1*) codifica uma proteína que interage com BRCA1, atuando no reparo de quebras dupla fita no DNA e nos processos relacionados à apoptose[35].

Através do estudo realizado por Couch *et al* (2017), pôde-se observar que variantes patogênicas em *BARD1* apresentaram um risco moderado para o desenvolvimento de câncer de mama [2,16 (IC 95%: 1,31-3,63)[86]. Porém, estudo publicado por Kurian e colaborados (2017), mostrou que pacientes com variantes patogênicas em *BARD1* apresentam um baixo risco para o desenvolvimento de câncer mama e de ovário [1,92 (IC 95%: 1,36 – 2,72) e 0,59 (IC 95%: 0,21 – 1,68), respectivamente][62]. Portanto, embora alguns estudos têm associado a presença de variantes germinativas em *BARD1* à um moderado risco para câncer de mama[86], mais estudos são necessários para confirmação, uma vez que variantes patogênicas em *BARD1* são raras (<1 em 500 em pacientes com câncer de mama).

Apesar de um dos estudos mencionados acima sugerir um risco aumentado de câncer de mama em portadores de variantes patogênicas em *BARD1,* as diretrizes do NCCN enfatizam que mais evidências para fornecer recomendações acerca do rastreio de câncer de mama são necessárias.

1.4.1.2.4. *CHEK2*

O gene *CHEK2* (do inglês: *checkpoint kinase 2*) codifica uma proteína do tipo quinase, determinando a parada do ciclo celular e fosforilando vários substratos proteicos (incluindo p53 e BRCA1) em resposta à danos no DNA[93].

A variante germinativa mais comumente encontrada em *CHEK2* é a c.1100delC (p.Thr367Metfs), a qual foi identificada e associada aos tumores de mama e colorretal[94]. Estudos publicados por diversos grupos têm demonstrado que a variante c.1100delC aumenta o risco de câncer de mama (RR: 2,7; IC 95%: 2,1-3,4) em indivíduos testados negativos para *BRCA1* e *BRCA2*[95–97]. Além disso, Cybulski e colaboradores, através da genotipagem de três variantes (c.444+1G>A, c.1100delC e c.470T>C), demonstrou um risco aumentado para múltiplos tipos de tumores, incluindo os cânceres de mama, colorretal, próstata e tireóide[98]. Ainda, outro estudo realizado por Cybulski *et al.* (2011) descreveu que portadores de variantes que levam à produção de uma proteína truncada apresentam um risco cumulativo de 20-44% para o desenvolvimento de câncer, dependendo da história familiar[99]. Adicionalmente, a presença de variantes germinativas em homozigose no gene *CHEK2* (c.1100delC) foi relatada por Adank e colaboradores (2013), sendo que esses indivíduos apresentaram um risco seis vezes maior de desenvolvimento de câncer de mama em relação aos heterozigotos[100]. Recentemente, Kurian *et al.* (2017) mostraram que portadores de variantes patogênicas em *CHEK2* apresentam um risco moderado para o desenvolvimento de câncer de mama [2,12 (IC 95%: 1,63-2,77) e um baixo risco [0,50 (IC 95%: 0,21-1,12)] para o desenvolvimento de câncer de ovário[62].

De acordo com as diretrizes do NCCN, para os portadores de variantes patogênicas em *CHEK2* deve ser recomendada a realização de mamografia anual a patir dos 40 anos de idade, levando em consideração a ressonância magnética de acordo com a história pessoal e familiar de câncer. Vale destacar que a mastectomia profilática pode ser considerada de acordo com a história familiar de câncer.

1.4.1.2.5. Parálogos de *RAD51*

Outros genes envolvidos no reparo de quebras dupla fita do DNA foram associados a um risco aumentado de câncer de mama e/ou ovário, entre eles os parálogos de *RAD51* (genes *RAD51B*[101,102], *RAD51C*[102,103] e *RAD51D*[102,104]).

O gene *RAD51B* tem recebido destaque em alguns trabalhos envolvendo tumores de mama. Polimorfismos no gene *RAD51B* foram associadas com aumento no risco de câncer de mama em estudos de associação gênomica (do inglês: *Genome-Wide Association Study – GWAS*). Os polimorfismos rs999737 (intron 10) e rs2588809 (intron 7) foram associados a um risco aumentado [RR: 1,27 (IC 95%: 1,09-1,48) e 0,98 (IC 95%: 0,82-1.16), respectivamente] de câncer de mama em mulheres[105,106]. Além disso, outro polimorfismo, também localizado no intron 7 do gene *RAD51B*, rs1314913, foi associado com o risco aumentado (OR: 1,54, IC 95%: 1,35-1,78) de desenvolvimento de câncer de mama em homens[107]. Poucas variantes germinativas que levam à perda de função em *RAD51B* estão descritas na literatura. Golmard e colaboradores (2013) descreveram duas variantes em pacientes diagnosticadas com câncer de mama: uma que afeta o *splicing* (c.452+3A>G) em uma família com história de câncer de mama e ovário, e a outra variante do tipo *missense*, predita como deletéria por ferramentas *in silico* (c.475C >T, p.Arg159Cys) em uma família com câncer de mama[108]. Além disso, Song *et al*. (2015) relataram a presença de uma variante *nonsense* (c.489T>G; p.Tyr163Ter) e de *splicing* (c.854-2A>G) em pacientes diagnosticadas com câncer de ovário, não selecionadas pela história familiar de câncer[102]. Já no estudo publicado por Li *et al*. (2018), foram identificadas duas portadoras com a mesma variante do tipo *nonsense* (c.139C>T, p.Arg47Ter). Em relação à história familiar de câncer, ambas relataram a presença de cânceres de mama e ovário, sendo que uma portadora também informou a presença de melanoma[109]. Devido à raridade das variantes com perda de função descritas na literatura e em bancos de dados envolvendo esse gene, o risco relativo para o desenvolvimento dos cânceres de mama e ovário permanecem inconclusivos.

O *RAD51C* é um gene fundamental na recombinação homóloga, sendo que variantes nesse gene foram relacionadas à Anemia de Fanconi[110]. O *RAD51C* foi pesquisado como um possível gene de predisposição aos tumores de mama e de ovário em 1.100 famílias de alto risco, as quais foram anteriormente testadas negativas em *BRCA1*/*BRCA2*. Variantes germinativas foram identificadas em 1,3% das famílias com câncer de mama e ovário. Porém, nenhuma variante patogênica foi identificada em famílias com apenas diagnóstico de câncer

de mama[111]. Adicionalmente, Pelttari e colaboradores observaram que variantes germinativas em *RAD51C* estavam presentes somente em famílias com diagnóstico de câncer de ovário[112]. No estudo recentemente publicado por Kurian e colaboradores (2017), os autores sugerem que variantes patogênicas em *RAD51C* não aumentam expressivamente o risco para o desenvolvimento de câncer de mama (RR: 1,43; IC 95%: 0,97-2,12)[62]. Vale destacar ainda que, segundo estudo realizado por Shimelis *et al.* (2018) variantes patogênicas em *RAD51C* foi associada a risco moderado (OR: 2,64; IC 95%: 1,44-4,80) para o desenvolvimento de câncer de mama triplo-negativo[60]. Em relação ao câncer de ovário, Kurian *et al.* (2017) reportou que o risco relativo para o desenvolvimento desse tipo de câncer de é de 4,98 (IC 95%: 3,09-8,04)[62]. Adicionalmente, estudo realizado por LaDuca *et al.* (2018) reportou que variantes patogênicas em *RAD51C* conferem um alto risco para o desenvolvimento de câncer de ovário (OR: 8,3; IC 95%: 5,43-12,48)[113]. Diante do exposto, a inclusão do gene *RAD51C* nos testes clínicos de rotina, para pacientes em risco para o desenvolvimento de câncer de mama, ainda é controversa, principalmente devido à baixa incidência ou ainda, à falta de identificação de variantes populações-específicas.

Em relação ao gene *RAD51D*, alguns estudos mostraram a presença de variantes germinativas patogênicas em pacientes diagnosticados com câncer de mama, selecionados devido à idade precoce ao diagnóstico e história familiar de câncer. Nestes trabalhos, a frequência de variantes patogênicas em *RAD51D* variou de 0,07% a 0,19%[86,114,115]. Recentemente, Konstanta e colaboradores (2018) investigaram a presença de variantes patogênicas em *RAD51D* em 609 mulheres diagnosticadas com câncer de ovário e 569 mulheres com câncer de mama, todas não portadoras de variantes patogênicas em *BRCA1*/*BRCA2*. Os autores identificaram quatro variantes patogênicas, sendo que, três variantes foram observadas em pacientes com câncer de ovário (frequência de 0,53%) e apenas uma variante foi identificada em uma paciente diagnosticada com câncer de mama (frequência de 0,16%). Além disso, foi destacado que nas quatro famílias portadoras de variantes em *RAD51D*, um total de 13 casos de câncer de mama e quatro casos de câncer de ovário foram observdos[116]. Outro estudo publicado em 2018 por Chen *et al.*, através da análise envolvendo 7.657 pacientes com câncer de mama e anteriormente testadas negativas em *BRCA1*/*BRCA2*, detectou 29 casos (0,38%) com variantes patogênicas em *RAD51D*. Os autores enfatizaram que a variante c.270_271dup foi observada em 18 pacientes (0,24%), sendo que a mesma é considerada patogênica pelo banco de dados ClinVar. Entretanto, a mesma

variante foi encontrada em oito controles saudáveis (0,1%, 8/7947). Ainda, a variante patogênica c.270_271dup foi significativamente associada a um aumento do risco de câncer de mama (OR: 2,34, IC 95%: 1,02-5,38) e foram mais propensos a ser do tipo molecular triplo-negativo (p=0,003)[117]. Em relação ao câncer de ovário, já foi relatado que a presença de variantes patogênicas em *RAD51D* demonstrou estar associada a um risco aumentado para o desenvolvimento desse tipo de tumor[102,116,118]. Estudo realizado por Loveday *et al.* (2012), envolvendo 911 pacientes e 1.060 controles, mostrou uma associação entre a presença de variantes patogênicas em *RAD51D* e o aumento do risco de câncer de ovário (RR: 6,30; IC 95%: 2,86-13,85)[118]. Em um outro estudo, através da análise de 3.429 mulheres com câncer de ovário epitelial e 2.772 controles, os autores reportaram um risco superior ao mencionado acima (OR: 12,0; IC 95%: 1,5-90,0) para as mulheres com variantes patogênicas em *RAD51D*[102].

Segundo as diretrizes do NCCN, é recomendado que a realização da salpingo-ooforectomia redutora de risco seja considerada para as portadoras de variantes patogênicas em *RAD51C* e *RAD51D*, a partir dos 45 anos de idade. Atualmente, não existem evidências suficientes de que as variantes patogênicas em *RAD51C* e *RAD51D* estejam associadas a um risco aumentado de câncer de mama, com isso, o NCCN recomenda a realização de mamagrafia a partir dos 40 anos de idade[42].

1.4.1.3. Genes de baixo risco

Vários *loci* de susceptibilidade ao câncer de mama e/ou ovário foram associados à um risco aumentado de desenvolvimento de câncer. Esses *loci* podem seguir o modelo poligênico, ou podem agir sinergicamente com fatores ambientais ou de estilo de vida. A maioria desses *loci* de baixa susceptibilidade foi detectada através de estudos de associação genômica (*GWAS*), sendo que a maioria das alterações identificadas estão localizadas nas regiões intrônicas e intergênicas.

Inicialmente, estudo realizado por Easton e colaboradores através da análise de *GWAS* envolvendo pacientes diagnosticados com câncer de mama e história familiar sugestiva de câncer hereditário, identificou aproximadamente 12 mil polimorfismos que apresentaram uma associação com câncer de mama[119]. Seguido à isso, os autores selecionaram 30 polimorfismos mais significativos para serem validados em 21.860 casos e 22.578 controles. Os autores observaram seis polimorfismos associados a um risco aumentado para o

desenvolvimento de câncer de mama, incluindo polimorfismos nos genes *FGFR2* (rs2981582), *LSP1* (rs3817198), *MAP3K1* (rs889312), *TOX3* (rs12443621 e rs8051542) e um polimorfismo na região do cromossomo 8q24 (rs13281615)[120]. Posteriormente, essas regiões gênicas foram comparadas de acordo com *status* do receptor de estrógeno (ER) em 23.039 casos e 26.273 controles. Com isso, foi identifcado que os polimorfismos rs2981582 (*FGFR2*) e rs13281615 (8q24) foram associados à positividade do ER[121].

Em relação ao câncer de ovário, Goode *et al.* realizou um estudo de *GWAS* em 1.768 mulheres com câncer de ovário e 2.354 controles. A identificação de um *locus* de susceptibilidade em 9p22 (*BNC2*) foi associado a um risco de câncer de ovário. Além disso, foram identificados nove *loci* candidatos após estratificação dos casos de acordo com a histologia. Os autores observaram que os polimorfismos presentes nos genes *HOXD1*, *MYC*, *TIPARP*, *SKAP1* e *BNC2* podem ter um papel adicional no desenvolvimento do câncer de ovário[122].

Embora a existência e o risco conferido por *loci* de susceptibilidade ao câncer de mama e/ou ovário hereditário de baixa penetrância seja discutível, é importante salientar que a identificação desses alelos poderia explicar um subconjunto de casos de câncer de mama[123]. Além disso, têm sido proposto um modelo poligênico que incorpora múltiplos genes de baixa penetrância que levaria ao desenvolvimento de câncer[124–126].

1.5. Painéis gênicos

Como mencionado acima, variantes germinativas patogênicas nos genes *BRCA1* e *BRCA2*, que predispõem à HBOC, explicam até 25% dos tumores de mama e ovário hereditários, ou seja, não respondem pela totalidade dos casos hereditários[28–30]. Por esse motivo, diversos trabalhos vêm enfatizando a utilização de painéis gênicos para investigação da presença de variantes patogênicas em outros genes que não *BRCA1* e *BRCA2* em famílias HBOC.

Um dos primeiros trabalhos envolvendo painéis gênicos em famílias com critérios para HBOC foi realizado em 2010 por Walsh e colaboradores. O painel gênico utilizado foi composto por 21 genes, sendo testado inicialmente em 20 mulheres com variantes germinativas conhecidas. A utilização dessa metodologia demonstrou ser altamente sensível, uma vez que todas variantes conhecidas foram identificadas[18]. Em seguida, o mesmo painel gênico foi

utilizado em 273 mulheres com câncer de ovário, 48 com câncer de peritônio, 31 com câncer de tuba uterina e 8 com carcinoma de endométrio. Todos os casos foram selecionados no momento do diagnóstico e não por idade ou histórico familiar. Dos 360 indivíduos, 24% eram portadores de variante germinativa associada à perda de função proteica, sendo 18% em *BRCA1* ou *BRCA2* e 6% em um dos seguintes genes: *BARD1*, *BRIP1*, *CHEK2*, *MRE11A*, *MSH6*, *NBN*, *PALB2*, *RAD50*, *RAD51C* ou *TP53*[18].

Recentemente, estudo publicado por Buyss e colaboradores, teve como objetivo avaliar os resultados de testes genéticos em mulheres que apresentavam história pessoal de câncer de mama. A população de estudo foi composta por 35.409 mulheres com diagnóstico de apenas um tumor de mama e que foram submetidas à realização de um painel gênico abrangendo 25 genes. Variantes patogênicas foram identificadas em 9,3% das mulheres testadas, sendo que 51,5% dessas variantes estavam presentes em outros genes que não *BRCA1*/*BRCA2*: 11,7% no gene *CHEK2,* 9,7% no gene *ATM* e 9,3% no gene *PALB2*. Além disso, os autores relataram que a prevalência de variantes patogênicas nos genes *BARD1*, *BRIP1* e *RAD51C* foi estatisticamente maior entre mulheres com tumores de mama do subtipo molecular triplo negativo[89].

Outro estudo publicado recentemente por Couch *et al.* (2017), envolveu a análise de 21 genes em 65.057 pacientes com câncer de mama. A frequência de variantes patogênicas identificada em 41.611 mulheres foi de 10,2%. Após a exclusão de genes já associados ao câncer de mama (*BRCA1*, *BRCA2, CDH1*, *PTEN* e *TP53*), foram observadas variantes patogênicas em 5 de 16 genes, destacando: *ATM* (RR: 2,78), *BARD1* (RR: 2,16), *CHEK2* (RR: 1,48), *PALB2* (RR: 7,46) e *RAD51D* (RR: 3,07). Ainda, variantes nos genes associados a um risco aumentado de câncer de ovário (*BRIP1*, *RAD51C, MLH1* e *PMS2*) não foram associados com o aumento do risco de desenvolvimento do câncer mama[86].

Com base nesses estudos, podemos observar que ao lado de genes de alto risco, como *BRCA1*/*BRCA2* e genes do complexo de reparo de *Mismatch Repair* (MMR), outros genes também contribuem para a predisposição familiar ao câncer e, sem dúvida, tecnologias de nova geração podem e devem ser utilizadas para identificar as causas genéticas entre as famílias que testaram negativo para variantes em genes de alto risco através de métodos tradicionais[115,127].

1.6. À procura pelo BRCAX

Embora a identificação e teste genético de famílias de alto risco para síndromes de predisposição hereditária podem levar à identificação dos fatores causais (alterações genéticas) associados ao desenvolvimento de múltiplos casos de tumores característicos das síndromes de predisposição hereditária ao câncer, é observado que uma parcela de indivíduos com diagnóstico clínico de determinada síndrome não possui variantes em genes clássicos conhecidos como associados com o fenótipo[128]. Para as famílias de alto risco para HBOC e sem variante germinativa identificada nos genes classicamente associados (como *BRCA1* e *BRCA2*), a hipótese primária está relacionada com a existência de um ou mais genes adicionais de alto risco (a exemplo de *CDH1/PALB2/PTEN/STK11/TP53*). No entanto, os estudos até agora realizados mostraram que, apesar de outros genes de alto e moderado risco também responderem por uma parcela dos casos de câncer de mama ou ovário hereditários, existe uma parcela significativa de casos cuja causa da predisposição hereditária permanece desconhecida[129], sugerindo que, alelos adicionais de alto risco deveriam existir, mas em uma frequência inferior àquela dos genes *BRCA1/BRCA2*. Dessa maneira estudos adicionais envolvendo análises mais detalhadas e aprofundadas fazem-se necessários.

1.6.1. Análises em larga escala

1.6.1.1. Alterações gênicas

Análises moleculares em larga escala estão contribuindo muito para a identificação de genes relacionados à doenças monogênicas ou ainda à doenças com alta heterogeneidade molecular, como no caso do estudo realizado por Lupski e colaboradores, que através de análises de larga escala encontraram um novo alelo no gene *SH3TC2* relacionado com a doença de Charcot-Marie-Tooth[130]. Dependendo da doença, entre 70-92% dos pacientes não são diagnosticados como portadores de variantes germinativas, mesmo após a realização do teste genético envolvendo painéis gênicos[131,132]. É possível que variantes em genes (ou regiões) não incluídos nos painéis possam estar contribuindo para um risco aumentado de se desenvolver o câncer.

Diversos pesquisadores apoiam a ideia de que análises completas do exoma (do inglês: *Whole-Exome Sequencing* – WES) ou ainda, do genoma (do inglês: *Whole-Genome Sequencing* – WGS) são as ferramentas mais apropriadas para identificar as alterações genéticas nas síndromes familiares. Conforme Snape e colaboradores, a análise de exomas oferece o

potencial de realizar uma estratégia independente de conhecimento prévio ("*agnostic approach*"), isto é, uma estratégia não baseada em gene candidato, mas que faça uma análise ao longo de todo o exoma à procura de possíveis variantes com potencial significado para a doença em questão[133].

Com a utilização crescente, os estudos de WES e WGS permitiram a identificação de novos genes considerados de alto e moderado risco em vários tipos tumorais, como os genes de susceptibilidade ao câncer de pâncreas (*PALB2* e *ATM*)[134], à feocromocitoma (*MAX*)[135] e ainda, os genes *POLD1* e *POLE*, os quais foram associados ao câncer colorretal hereditário[136]. Além disso, diversos estudos reportaram a presença de variantes em genes que aumentam o risco de desenvolvimento de câncer de mama e/ou ovário como *FANCC*[137,138] e *BLM*[138]. Thompson e colaboradores identificaram, através de análise de exoma, novas variantes patogênicas nos genes de reparo *FANCC* e *BLM* como potenciais alelos causadores de aumento da suscetibilidade ao câncer de mama[138]. Ainda, Kiiski e colaboradores realizaram a análise do exoma de 24 pacientes com câncer de mama, provenientes de 11 famílias finlandesas com câncer de mama. Dentre as alterações encontradas no gene *FANCM*, a variante do tipo *nonsense* c.5101C>T (p. Gln1701Ter) foi significativamente mais frequente entre os pacientes com câncer de mama do que entre os controles (RR: 1,86; IC:95%: 1,26-2,75). Além disso, quando analisados apenas os tumores de mama triplo-negativos foi observada uma maior frequência da variante relatada (RR: 3,56; IC:95%: 1,81-6,98). Com isso, os autores sugerem que o gene *FANCM* atuaria como um gene de susceptibilidade ao câncer de mama, conferindo uma maior predisposição para o desenvolvimento do câncer de mama triplo-negativo[139].

1.6.1.2. Alterações no número de cópias cromossômicas

Além da identificação de variantes pontuais no DNA (SNPs), as análises em larga escala têm ajudado a identificar alterações numéricas de cromossomos, alelos ou genes, alterações essas que podem ter impactos severos na expressão e função de diversos genes[140]. Uma das técnicas amplamente utilizada para a análise de ganhos e perdas cromossômicas é a hibridização genômica comparativa por *microarray* (aCGH). Esta técnica apresenta alta resolução e possibilita o mapeamento direto de aberrações por todo o genoma humano[141].

A pesquisa de alterações cromossômicas numéricas tem permitido avançar nosso conhecimento em relação à biologia do câncer, através da identificação de aberrações genéticas subjacentes à iniciação, desenvolvimento e metástases[142]. De maneira adicional, esse tipo de análise pode levar à identificação de novos genes associados ao desenvolvimento de tumores. Estudo realizado por Sabatier e colaboradores em 2010, permitiu a identificação de alterações em homozigose no gene *BARD1*. Os autores descreveram o caso de uma mulher diagnosticada com câncer de mama aos 36 anos e história familiar positiva de câncer. Através das análises do tumor, foi observado a inativação de BRCA1 devido à deleção de *BARD1*[143]. Além disso, a deleção de *BARD1* foi associada à várias características dos tumores associadas a *BRCAness* (tumores que apresentam um defeito no reparo do DNA por recombinação homóloga, com características semelhantes aos tumores que com perda de BRCA1/BRCA2[144]), tais como idade precoce ao diagnóstico, tipo ductal, alto grau histológico e triplo negatividade[145].

No que se refere às famílias de alto risco para câncer de mama e/ou ovário hereditários, evidências da literatura têm apontado que tumores agrupados como BRCAX (dado o desconhecimento do gene associado ao seu desenvolvimento) apresentam uma alta heterogeneidade genética, molecular e histopatológica[146,147]. Uma análise mais detalhada do perfil de alterações não somente gênicas, mas também cromossômicas pode auxiliar na identificação de regiões-alvo para a pesquisa de oncogenes e genes supressores tumorais, os quais poderiam estar diretamente envolvidos na iniciação e progressão tumoral. A partir das análises já realizadas e disponíveis na literatura, envolvendo pacientes classificados como BRCAX, foi possível identificar várias regiões cromossômicas que podem estar associadas a outros genes de predisposição ao câncer de mama e/ou ovário[148–151].

Estudo publicado por Gronwald *et al.* (2005), envolveu a análise de alterações cromossomais em material tumoral de pacientes com câncer de mama classificados como BRCAX *versus* tumores de mama esporádicos. Dentre os achados, os autores destacaram que a perda de 8p e ganhos em 8q, 19q, bem como ganhos de 20q (em tumores do tipo ductal) foram detectados mais frequentemente em BRCAX do que nos cânceres esporádicos. A análise da perda de 8p e do ganho de 8q mostrou que estas alterações poderiam se constituir em eventos iniciais na tumorigênese dos tumores BRCAX[148].

De maneira similar, Didraga e colaboradores, através da tecnologia de aCGH, identificaram diferentes alterações no número de cópias gênicas em tumores de mama

analisados. Dentre os tumores estudados, 28 eram provenientes de pacientes com variante germinativa em *BRCA1*, 28 amostras de pacientes com variante germinativa em *BRCA2*, 58 pacientes sem variante germinativa identificada, mas com história familiar positiva (BRCAX) e 49 tumores esporádicos. Com base nos resultados fornecidos, podemos destacar o ganho de quase todo o cromossomo 22 nos pacientes BRCAX. Outras alterações foram identificadas (ganho de 1q e 8q, perda de 8p e 11q) em vários casos de câncer de mama provenientes de famílias BRCAX[150]. Posteriormente, o mesmo grupo de pesquisa realizou a análise de segregação em seis famílias provenientes das pacientes que apresentaram alteração no cromossomo 22. Através da análise de ligação, foi observado que a região no cromossomo 4 apresentou um LOD *score* de 2,49, indicando a evidência de ligação. Em seguida, foi analisado o DNA germinativo de dois pacientes de cada família através do sequenciamento exômico. Inicialmente a análise foi restrita ao cromossomo 4, porém não foram identificadas variantes potencialmente patogênicas em mais de uma família. Com isso, todas as variantes foram re-analisadas, e foram detectadas inúmeras variantes classificadas como possivelmente patogênicas em genes que codificam proteínas responsáveis pela manutenção da integridade do DNA[152].

Apesar de alguns estudos (como por exemplo os acima citados) apontarem para algumas regiões cromossômicas potencialmente envolvidas com o surgimento e progressão dos tumores de mama classificados como BRCAX, esses achados precisam ser confirmados por estudos mais extensos e em outras populações, pois apenas um número limitado de casos foi analisado e há relativamente poucos relatos na literatura.

2. JUSTIFICATIVA

Apesar dos avanços nas técnicas de biologia molecular, os genes até agora associados ao câncer de mama e/ou ovário hereditário, sejam eles de alto, moderado ou baixo risco, respondem por aproximadamente de 50% dos casos. Determinar o risco de câncer de uma forma mais acurada para a grande proporção das mulheres com história pessoal e familiar de câncer de mama fortemente sugestiva de HBOC e que não possuem variante germinativa nos genes classicamente associados a esse tipo de neoplasia é fundamental. O conhecimento de qual gene está alterado e qual o risco associado a essa alteração possibilita uma melhora significativa nas decisões acerca do manejo do risco, bem como pode aumentar a gama de estratégias preventivas e de redução de risco (exemplo cirurgias profiláticas) a serem oferecidas.

Dessa forma, considerando (*i*) o alto risco de desenvolvimento de câncer associado à presença de variantes germinativas, (*ii*) o fato de que essas alterações respondem por apenas 50% dos casos de câncer de mama/ou ovário hereditário, bem como (*iii*) a possibilidade de que alterações gênicas ou cromossomais possam afetar o risco de desenvolvimento do câncer, acreditamos que a combinação de metodologias amplas e robustas seja necessária para o diagnóstico molecular do câncer de mama e/ou ovário hereditário e para a descoberta de novos mecanismos moleculares causais ou relacionados ao desenvolvimento desse tipo de neoplasia.

3. OBJETIVOS

3.1. Objetivo Geral

Identificar alterações genéticas em um grupo de mulheres com história pessoal e familiar sugestiva de câncer de mama e/ou ovário hereditário, negativas para variantes germinativas patogênicas nos genes *BRCA1* e *BRCA2*.

3.2. Objetivos Específicos

I) Verificar a presença e a frequência da variante germinativa fundadora portuguesa no gene *BRCA2* (c.156_157insAlu) em famílias de alto risco para HBOC;

II) Identificar alterações no número de cópias cromossômicas em amostras tumorais de mulheres com história pessoal e familiar sugestiva de câncer de mama e/ou ovário hereditário, através de aCGH;

III) Identificar variantes germinativas patogênicas/possivelmente patogênicas nas amostras *BRCA1/BRCA2*-negativas através de sequenciamento exômico.

4. ESTUDO DA PRESENÇA DE VARIANTE GERMINATIVA FUNDADORA PORTUGUESA NO GENE *BRCA2* (c.156_157insAlu), ATRAVÉS DA PCR SEGUIDA DO SEQUENCIAMENTO DE SANGER

Os resultados apresentados nessa seção foram publicados no periódico "*Cancer Genetics*" – fator de impacto: 2,183. O arquivo PDF do artigo publicado na revista científica encontra-se anexado à esta tese (Anexo I).

4.1. Screening and characterization of *BRCA2* c.156_157insAlu in Brazil: results from 1380 individuals from South and Southeast.

Felicio PS[§], Alemar B[§], Coelho AS, Berardinelli GN, Melendez ME, Lengert AVH, Michelli RD, Reis RM, Fernandes GC, Ewald IP, Bittar CM, Netto CBO, Artigalas O, Peixoto A, Pinheiro M, Teixeira MR, Vargas FR, Dos Santos ACE, Moreira MAM, Ashton-Prolla P[#], Palmero EI[#].

[§] These authors contributed equally to this work

[#] These authors jointly supervised this work

Cancer Genet. 2018 Dec;228-229:93-97.

Highlights

Portuguese immigration to Brazil greatly contributed to the genetic composition of current Brazilian population. In this study, we evaluated the frequency of a Portuguese founder Alu insertion in *BRCA2* exon 3 (c.156_157insAlu) among individuals fulfilling HBOC criteria in 1,380 unrelated Brazilian individuals, identifying the mutation in nine (0.65%) families. In probands with the *BRCA2* rearrangement, the European ancestry proportion was the most frequent (80%), followed by the African (10%). Although the haplotype was not informative for three families, the remaining six presented a haplotype compatible with the Portuguese ancestral haplotype. The relatively low mutational prevalence found in this cohort does not rule out the importance of testing for this mutation in at-risk individuals and the absence of routine testing for this mutation in the past years may have led to an underestimation of its frequency in our population.

Abstract

Portuguese immigration to Brazil occurred in several waves and greatly contributed to the genetic composition of current Brazilian population. In this study, we evaluated the frequency of a Portuguese founder Alu insertion in *BRCA2* exon 3 (c.156_157insAlu) among individuals fulfilling Hereditary Breast and Ovarian Cancer (HBOC) syndrome criteria in 1,380 unrelated families originated from three distinct Brazilian States. We identified the c.156_157insAlu *BRCA2* mutation in nine (9/1,380; 0.65%) probands analyzed. In carrier probands, European ancestry had the highest proportion (80%), followed by the African (10%) and Amerindian and in most families with the rearrangement, haplotype analyses were compatible with the Portuguese ancestral haplotype. In conclusion, the present study reports a low albeit relevant frequency of the Portuguese *BRCA2* founder mutation c.156_157insAlu in Brazilian patients at-risk for HBOC Brazilian population.

Key words: *BRCA2*, HBOC, genetic screening, Alu elements

Introduction

Alu elements are retroelements (retrotransposons) typically 300nt in size, which account for approximately 11% of the human genome. It is believed that genomic rearrangements induced by Alu elements account for approximately 0.1% of human disease [1, 2]. Indeed, pathogenic Alu insertions have been described in many cancer risk genes, such as *APC*, ATM*, CHEK2*, *BRCA1* and *BRCA2* [3]. In 2005, Teugels *et al.* described an Alu insertion within exon 3 of *BRCA2* (c.156_157insAlu) in a female breast cancer patient of Portuguese origin fulfilling clinical criteria of Hereditary Breast and Ovarian Cancer (HBOC) syndrome [4]. Subsequent analysis showed that the insertion co-segregated with the cancer phenotype in that family, and resulted in exon 3 skipping, a region that contains an important *BRCA2* regulatory domain [4]. Further analysis has shown that *BRCA2* c.156_157insAlu is a Portuguese founder mutation accounting for 27-38% of all *BRCA1* and *BRCA2* pathogenic mutations in HBOC families originating mostly from northern/central Portugal [5, 6]. Moreover, the estimated cumulative incidence of breast cancer in women carrying this mutation is at least as high as that among carriers of other *BRCA2* mutations [7].

Since Brazil was a Portuguese colony and received more than 2.2 million of Portuguese immigrants between 1500 and 1991 [8], in this study we sought to evaluate the frequency of this mutation in Brazilian individuals fulfilling the HBOC criteria. Furthermore, we performed ancestry and haplotype analyses of the mutation carriers to gain insight into the ancestral origin of the *BRCA2* c.156_157insAlu.

Materials and Methods

Ethical aspects and patient recruitment

Unrelated individuals fulfilling HBOC criteria were recruited from three distinct Brazilian States, at three reference centers for Genetic Cancer Risk Assessment related to the Brazilian Hereditary Cancer Network (BHCN): Barretos Cancer Hospital (BCH, Barretos/São Paulo), Hospital de Clínicas de Porto Alegre (HCPA, Porto Alegre/Rio Grande do Sul) and Instituto Nacional de Câncer (INCA, Rio de Janeiro/Rio de Janeiro). Although all participants had a personal and/or family history suggestive of HBOC syndrome, each institution used slightly distinct testing criteria, detailed in Supplementary Methods and in Palmero *et al.* [9]. Of note, all patients recruited for this study fulfilled at least one of the NCCN HBOC genetic testing criteria. Ethical approval for this study was obtained from the institutional ethics committees of all participating centers and all patients were offered genetic counseling.

DNA isolation and *BRCA2* c.156_157insAlu genotyping

Genomic DNA was isolated from peripheral leucocytes using the Qiagen DNeasy kit or the Qiagen Flexigene kit following the manufacturer's instructions. *BRCA2* c.156_157insAlu genotyping was performed as previously described, using two distinct PCR reactions [5]. Positive and negative controls were used in all experiments and all positive cases were confirmed in a second independent sample.

Ancestry analyses

A panel of 46 AIM-INDELs was used and analyses were done in a single multiplex PCR followed by capillary electrophoresis, as previously described [10]. The electropherograms were analyzed and genotypes were automatically assigned with GeneMapper v4.1 (Applied Biosystems). Ancestry proportions were obtained using the Structure v2.3.3 software [11, 12], taking into account the four main population groups: European (EUR), African (AFR), Native

American (NAM) and East Asian (EAS). Ancestry was analyzed in the nine probands and 18 relatives with the *BRCA2* c.156_157insAlu variant.

Haplotype analyses

For haplotype analyses we assayed nine microsatellite markers flanking the *BRCA2* gene by PCR using fluorescently-labeled primers. PCR products were run on an ABI PRISM 310 Genetic Analyzer (Applied Biosystems) with the fluorescence labeled DNA fragment size standard TAMRA. Haplotype construction was performed manually based on the genotypes obtained of probands and their relatives [5]. Haplotype analyses were performed in the nine probands and 18 relatives with *BRCA2* c.156_157insAlu variant.

Results

Screening for the *BRCA2* c.156_157insAlu was performed in 1,380 apparently unrelated probands referred to Genetic Cancer Risk Assessment at Barretos (São Paulo State, n=696), Porto Alegre (Rio Grande do Sul State, n=376) and Rio de Janeiro (Rio de Janeiro State, n=308), due to a suspicion of HBOC syndrome. Nine carriers were identified, corresponding to an overall mutation prevalence of 0.65%. Three carriers were referred from São Paulo State, two from Rio Grande do Sul, and four from Rio de Janeiro, resulting in center-specific prevalence of 0.43%, 0.53% and 1.30%, respectively. Among carriers, 89% had breast cancer (8/9), with a mean age at diagnosis of 37.9 years. Only one proband had ovarian cancer, at 62 years old. Through genetic counseling and cascade testing of all available relatives, we identified 15 additional carriers in seven families.

Ancestry was analyzed in the nine probands and 18 relatives to estimate proportions of European (EUR), African (AFR), Amerindian (NAM) and East Asian (EAS) origin. All cases showed high proportion of European ancestry (range 38-94%), and the average ancestry proportions were 75.3% for EUR, 12.9% for AFR, 6.5% for NAM and 5.3% for EAS. Clinical, molecular and ancestry data are detailed in Table 1.

Table 1 - Detailed family history of the nine families carrying the *BRCA2* c.156_157insAlu

Individual	Molecular and clinical characterization			Ancestry results			
	c.156_157insAlu status	Kinship	Diagnosis (age, years)	African (%)	European (%)	East Asian (%)	Amerindian (%)
INCA-1*	Carrier	Proband	BC (38)	3,3	93,6	2,2	0,9
INCA-1b	Non-carrier	Proband's mother	Unaffected	1,9	93,8	2,5	1,9
INCA-1c	Carrier	Proband's father	Gastric Ca (68)	5,3	90,9	2,7	1,1
INCA-2*	Carrier	Proband	BC (48)	28,7	64,3	4,4	2,6
INCA-2a	Non-carrier	Proband's niece	Unaffected	34,0	60,5	2,2	3,4
INCA-2b	Non-carrier	Proband's niece	Unaffected	27,5	67,0	3,3	2,2
INCA-2c	Carrier	Proband's daughter	Unaffected	17,0	64,1	2,3	16,5
INCA-2d	Carrier	Proband's sister	Unaffected	25,8	71,2	1,6	1,4
INCA-3*	Carrier	Proband	BC (33)	15,0	75,5	3,0	6,4
INCA-3a	Carrier	Proband's mother	BC (42)	14,2	70,0	5,2	10,7
HCPA-1	Carrier	Proband	Bilat BC (51, 55)	2,3	84,2	3,7	9,8
HCPA-2	Carrier	Proband	BC (27)	10,2	84,6	3,1	2,1
HCB -1*	Carrier	Proband	BC (39)	2,6	92,2	1,7	3,5
HCB -1a	Carrier	Proband's father	Unaffected	2,6	80,2	12,5	4,6
HCB-2*	Carrier	Proband	BC (39)	23,2	64,9	2,2	9,7
HCB-2a	Carrier	Proband's sister	Unaffected	4,2	90,7	2,8	2,3
HCB-2b	Carrier	Proband's sister	BC (62)	19,3	63,7	7,4	9,7
HCB-2c	Carrier	Proband's niece	BC (28)	4,6	85,8	6,2	3,4
HCB-2d	Carrier	Proband's sister	BC (62)	2,7	61,6	28,0	7,6
HCB-2e	Carrier	Proband's sister	Unaffected	4,6	91,4	2,3	1,7
HCB-2f	Carrier	Proband's son	Unaffected	21,5	37,7	16,8	24,1
HCB-2g	Carrier	Proband's brother	Unaffected	3,4	90,4	3,4	2,8
HCB-3*	Carrier	Proband	Gastric cancer (62), OC (62)	3,6	84,6	9,4	2,5
HCB-3a	Carrier	Proband's son	Unaffected	33,9	52,3	3,2	10,6
HCB-3b	Carrier	Proband's sister	Unaffected	19,2	52,9	1,7	26,1
HCB-4*	Carrier	Proband	BC (29)	4,0	84,0	5,0	5,0
HCB-4a	Carrier	Proband's father	Kidney Ca (47)	13,0	81,5	4,0	1,4

Subtitles: HCB: Hospital do Câncer de Barretos (São Paulo State); INCA: Instituto Nacional de Câncer (Rio de Janeiro State), HCPA: Hospital de Clínicas de Porto Alegre (Rio Grande do Sul State), BC: breast cancer, OC: ovarian cancer, Ca: cancer, Bilat: bilateral

To evaluate the contribution of the founder Portuguese haplotype, all nine probands and their 18 relatives were genotyped for polymorphic microsatellite markers flanking *BRCA2*. Haplotypes were completely phased in three families (INCA-1, INCA-2 and HCB-2), differing from the ancestral haplotype only by one recombinational event (considering the most parsimonious relationships between these haplotypes). Additionally, the HCB-3 family,

although not completely phased, also differed from the ancestral haplotype by one recombinational event. Although two other families (HCB-1 and HBC-4) were not completely phased, their haplotypes were compatible with the Portuguese ancestral haplotype (Table 2). In the remaining families (HCPA-1, HCPA-2 and INCA-3) it was not possible to infer the haplotype with the data available.

Table 2 - Microsatellite markers genotypes in nine probands carrying the *BRCA2* c.156_157insAlu.

Markers	Consensus pattern	INCA-1	INCA-2	INCA-3	HCPA-1	HCPA-2	HCB-1	HCB-2	HCB-3	HBC-4	Portuguese ancestral haplotype H1
Centromere											
D13S1700	AGAA	**317**	253	**317**	249/305	329/333	**317**	249	**317**	**317**	**317**
D13S260	TG	**160**	162	**160**	154/168	160/162	**160**	162	**160**	**160**	**160**
D13S1698	GT	**156**	166	156/160	158/160	158/172	**156**	166	**156**	**156**	**156**
BRCA2											
D13S1701	TTCC	**299**	**299**	**299**	295/299	**299**	291/299	**299**	**299**	**299**	**299**
D13S171	TG	**230**	**230**	230/240	**230**	**230**	**230**	**230**	**230**	230/240	**230**
D13S1695	AC	**242**	**242**	258	242/252	242/250	**242**	**242**	252	242/250	**242**
D13S1694	TG	**228**	**228**	**228**	224/228	228/234	**228**	**228**	234	**228**	**228**
D13S310	GT	**144**	**144**	146	144/148	**144**	**144**	**144**	144/148	**144**	**144**
D13S267	TG	150	**158**	156	156/158	144/158	**158**	**158**	150	**158**	**158**
Telomere											

Markers present in the ancestral haplotype are highlighted in bold.

Discussion

Many current genetic testing methods applied to hereditary cancer cannot reliably detect large deletions and insertions related to retroelements, due to technical limitations [3]. In the past years, the two main diagnostic methods used in *BRCA* mutation screening were Sanger sequencing and Multiplex Ligation-dependent Probe Amplification (MLPA), and none of these methods were able to detect the c.156_156insAlu, which required a specific PCR-based test. While conventional PCR is still not capable of detecting such a large insertion as the *BRCA2* c.156_157insAlu, only recently (July 2016) a commercial MLPA kit for *BRCA2* was modified to include a specific probe for this mutation. Moreover, even considering the incorporation of next generation sequencing (NGS) into clinical genetic testing for hereditary cancer in the past few years, it is well known that the detection of large genomic

rearrangements by NGS platforms is only possible after extensive optimization and validation, and requires a specific platform and bioinformatics pipeline which are not always available in commercial laboratories.

In this study, we found nine carriers of c.156_157insAlu among 1,380 unrelated probands, corresponding to an overall mutation prevalence of 0.65%. From previous studies of Brazilian and non-Brazilian cohorts, we expected that nearly 20% of all individuals fulfilling standard (i.e. NCCN) criteria for HBOC would carry a pathogenic mutation in *BRCA1* or *BRCA2* [13-16]. Thus, in a cohort of 1,380 individuals we would estimate the presence of approximately 276 *BRCA* mutation carriers. In this scenario, the nine proband carriers of the *BRCA2* c.156_157insAlu found in our study would represent around 3.3% of all detected mutations and about one tenth of the proportion of this mutation among all *BRCA* mutations observed in Portugal [5, 6]. The lower mutational prevalence identified in our study is likely due to the high admixture observed in the Brazilian population, with genetic contributions from many distinct regions worldwide [17]. This is the consequence of five centuries of immigration and interethnic crosses of inhabitants from three areas: the European colonizers (mostly represented by the Portuguese), African slaves, and Amerindians. Although our ancestry analysis has shown a high proportion of European ancestry among all mutation carriers (75.3%), a recent study shows that individuals from the Southeast/South regions of Brazil derive from a wide European region, including central and northern Europe as well as parts of the Middle East. This profile is different from individuals originating in the northeast region of Brazil, which are expected to have a European ancestry mainly restricted to the Iberian Peninsula [18].

Among the nine proband carriers reported here, five were previously identified in smaller screening studies [13, 19, 20], and haplotype analysis was performed in three families without, however, a definitive conclusion about their ancestral haplotype [19]. Here, we performed haplotype reconstruction with nine microsatellites markers. Although the haplotype was not informative for three families, the remaining six presented haplotypes compatible with the ancestral Portuguese haplotype. These results suggest that these families share a common ancestor with the Portuguese *BRCA2* c.156_157insAlu positive families.

Despite the large number of Brazilian cases analyzed, our study is limited to the Southeast/South Brazilian regions. Thus, analysis in a larger cohort including individuals from all Brazilian regions would help to understand the overall contribution of *BRCA2*

c.156_157insAlu in our population, especially in regions with high Portuguese ancestry. Moreover, the relatively low mutational prevalence found in this cohort does not rule out the importance of testing for this mutation in at-risk individuals and the absence of routine testing for this mutation in the past years may have led to an underestimation of its frequency in our population.

Although not as common as in Portugal, the c.156_157insAlu is the third most frequently reported *BRCA2* mutation in Brazil, corresponding to more than one third of all large genomic rearrangements reported in the Brazilian population (Palmero *et al*., unpublished results), highlighting the importance of its identification.

Acknowledgments

We thank Gustavo Stumpf for his laboratory support and the Center of Molecular Diagnosis and the Oncogenetics Department from Barretos Cancer Hospital for their support.

Funding

This study was supported in part by grants from CNPq (408313/2016-1), FAPERGS (PRONEX: 16/2551-0000486-2), Hospital de Clínicas de Porto Alegre (FIPE-HCPA), FINEP-CT-INFRA (02/2010) and FAPESP (2013/24633-2). RMR, EIP and PA-P are recipients of National Council of Technological and Scientific Development (CNPq) scholarships.

References

[1] Zhang W, Edwards A, Fan W, Deininger P, Zhang K. Alu distribution and mutation types of cancer genes. BMC Genomics 2011;12:157.

[2] Deininger PL, Batzer MA. Alu repeats and human disease. Mol Genet Metab 1999;67:183-93.

[3] Qian Y, Mancini-DiNardo D, Judkins T, Cox HC, Brown K, Elias M, et al. Identification of pathogenic retrotransposon insertions in cancer predisposition genes. Cancer Genet 2017;216-217:159-69.

[4] Teugels E, De Brakeleer S, Goelen G, Lissens W, Sermijn E, De Grève J. De novo Alu element insertions targeted to a sequence common to the BRCA1 and BRCA2 genes. Hum Mutat 2005;26:284.

[5] Peixoto A, Santos C, Rocha P, Pinheiro M, Príncipe S, Pereira D, et al. The c.156_157insAlu BRCA2 rearrangement accounts for more than one-fourth of deleterious BRCA mutations in northern/central Portugal. Breast Cancer Res Treat 2009;114:31-8.

[6] Peixoto A, Santos C, Pinheiro M, Pinto P, Soares MJ, Rocha P, et al. International distribution and age estimation of the Portuguese BRCA2 c.156_157insAlu founder mutation. Breast Cancer Res Treat 2011;127:671-9.

[7] Mavaddat N, Peock S, Frost D, Ellis S, Platte R, Fineberg E, et al. Cancer risks for BRCA1 and BRCA2 mutation carriers: results from prospective analysis of EMBRACE. J Natl Cancer Inst 2013;105:812-22.

[8] IBGE. Brasil: 500 anos de povoamento. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística (IBGE), 2007.

[9] Palmero EI, Galvão HC, Fernandes GC, Paula AE, Oliveira JC, Souza CP, et al. Oncogenetics service and the Brazilian public health system: the experience of a reference Cancer Hospital. Genet Mol Biol 2016;39:168-77.

[10] Pereira R, Phillips C, Pinto N, Santos C, dos Santos SE, Amorim A, et al. Straightforward inference of ancestry and admixture proportions through ancestryinformative insertion deletion multiplexing. PLoS One 2012;7:e29684.

[11] Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics 2000;155:945-59.

[12] Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 2003;164:1567-87.

[13] Fernandes GC, Michelli RA, Galvão HC, Paula AE, Pereira R, Andrade CE, et al. Prevalence of BRCA1/BRCA2 mutations in a Brazilian population sample at-risk for hereditary breast cancer and characterization of its genetic ancestry. Oncotarget 2016.

[14] Kast K, Rhiem K, Wappenschmidt B, Hahnen E, Hauke J, Bluemcke B, et al. Prevalence of BRCA1/2 germline mutations in 21 401 families with breast and ovarian cancer. J Med Genet 2016;53:465-71.

[15] Alemar B, Gregório C, Herzog J, Matzenbacher Bittar C, Brinckmann Oliveira Netto C, Artigalas O, et al. BRCA1 and BRCA2 mutational profile and prevalence in hereditary breast and ovarian cancer (HBOC) probands from Southern Brazil: Are international testing criteria appropriate for this specific population? PLoS One 2017;12:e0187630.

[16] Alemar B, Herzog J, Brinckmann Oliveira Netto C, Artigalás O, Schwartz IV, Matzenbacher Bittar C, et al. Prevalence of Hispanic BRCA1 and BRCA2 mutations among hereditary breast and ovarian cancer patients from Brazil reveals differences among Latin American populations. Cancer Genet 2016;209:417-22.

[17] Ruiz-Linares A, Adhikari K, Acuña-Alonzo V, Quinto-Sanchez M, Jaramillo C, Arias W, et al. Admixture in Latin America: geographic structure, phenotypic diversity and self-perception of ancestry based on 7,342 individuals. PLoS Genet 2014;10:e1004572.

[18] Kehdy FS, Gouveia MH, Machado M, Magalhães WC, Horimoto AR, Horta BL, et al. Origin and dynamics of admixture in Brazilians and its effect on the pattern of deleterious mutations. Proc Natl Acad Sci U S A 2015;112:8696-701.

[19] Ewald IP, Cossio SL, Palmero EI, Pinheiro M, Nascimento IL, Machado TM, et al. BRCA1 and BRCA2 rearrangements in Brazilian individuals with Hereditary Breast and Ovarian Cancer Syndrome. Genet Mol Biol 2016;39:223-31.

[20] Moreira MA, Bobrovnitchaia IG, Lima MA, Santos AC, Ramos JP, Souza KR, et al. Portuguese c.156_157insAlu BRCA2 founder mutation: gastrointestinal and tongue neoplasias may be part of the phenotype. Fam Cancer 2012;11:657-60.

5. ESTUDO DE ALTERAÇÕES NO NÚMERO DE CÓPIAS CROMOSSÔMICAS EM AMOSTRAS TUMORAIS DE MULHERES COM HISTÓRIA PESSOAL E FAMILIAR SUGESTIVA DE CÂNCER DE MAMA HEREDITÁRIO, ATRAVÉS DE aCGH

Os resultados apresentados nessa seção foram publicados no periódico "*Oncotarget*", com acesso livre e fator de impacto: 5,168 (2016). O arquivo PDF do artigo publicado na revista científica encontra-se anexado à esta tese (Anexo II).

5.1. Genetic alterations detected by comparative genomic hybridization in BRCAX breast and ovarian cancers of Brazilian population.

Felicio PS, Bidinotto LT, Melendez ME, Grasel RS, Campacci N, Galvão HCR, Scapulatempo-Neto C, Dufloth RM, Evangelista AF, Palmero EI.

Oncotarget. 2018 Jun 8;9(44):27525-27534.

ABSTRACT

Background: About 5–10% of breast/ovarian cancers are hereditary. However, for a large proportion of cases (around 50%), the genetic cause remains unknown. These cases are grouped in a separated BRCAX category. The aim of this study was to identify genomic alterations in BRCA1/BRCA2 wild-type tumor samples from women with family history strongly suggestive of hereditary breast/ovarian cancer. **Results:** A cohort of 31 Brazilian women was included in the study. Using the GISTIC algorithm, we identified 20 regions with genomic gains and 31 with losses. The most frequent altered regions were 1q21.2, 6p22.1 and 8p23.3 in breast tumors and Xq26 and Xp22.32-22.31 among the ovarian cancer cases. An interesting association identified was the loss of 22q13.31-13.32 and the presence of ovarian cancer cases. Among the genes present in the frequently altered regions, we found *FGFR1*, *NSMCE2*, *CTTN*, *CRLF2*, *ERBB2*, *STARD3*, *MIR3201* and several genes of *RAET* and *ULBP* family. Conclusions: In conclusion, our results suggest that alterations on chromosomes 1, 6, 8 and X are common on BRCAX tumors and that the loss on 22q can be associated with the presence of ovarian cancer. **Methods:** DNA copy number alterations were analyzed by 60K array comparative genomic hybridization in breast and ovarian FFPE tumors.

INTRODUCTION

According to the World Health Organization, breast cancer (BC) is the most common tumor in women worldwide [1]. It is known that 5–10% of BC cases have a hereditary component [2], being characterized by the presence of germline mutations in the *BRCA1* [3] or *BRCA2* [4] genes, which are associated with the hereditary breast and ovarian cancer predisposition syndrome (HBOC). HBOC patients have strong personal and family histories of cancer. Moreover, these patients are characterized by early age-at-diagnosis of cancer, increased frequency of bilateral tumors, and two or more generations affected by cancer [5, 6].

Recent studies have shown that alterations in other susceptibility genes, mainly involved in the homologous recombination and DNA repair pathways, can be causal factors of hereditary breast and ovarian cancers [7]. In spite of that, the predisposing genetic cause of about 50% of the families at-risk for hereditary breast and ovarian cancers remains unknown [8, 9]. These families are grouped in a category called BRCAX.

Evidences from the literature have shown that BRCAX tumors are rather heterogeneous, involving several different histopathological subgroups and genetic alterations [10, 11]. Several authors have shown the presence of new high penetrance genes associated with breast and ovarian cancers [11–17]. However, the opinion of the scientific community is controversial. There are authors who argue that the incidence of BRCAX tumors is associated with rare syndromes in which BC is only one component [12, 15, 16]. Other authors believe that this type of tumor results from mutations in several genes with low penetrance or population-specific [11, 13, 14, 18].

Studies using array-comparative genomic hybridization (aCGH) technique suggest that several chromosomal regions are associated with the development of hereditary BC, highlighting gains at chromosomes 1q, 8q, 17q and 20q and losses within chromosomes 8p, 11q, 13q and 17p [19–24]. Despite these findings, more studies are necessary to a better understanding of BRCAX molecular events in hereditary breast cancer. In this regard, the aim of this study was to identify chromosomal and subchromosomal copy number alterations in tumor samples from Brazilian women without *BRCA1/BRCA2* germline mutations with family history strongly suggestive of HBOC syndrome.

MATERIALS AND METHODS

Ethics statement

All participants gave their consent to use tumor samples for academic genetic research. In addition, the ethics committee of the Barretos Cancer Hospital (BCH) approved this study (approval number: 40814115.4.0000.5437).

Patients

This study included 31 unrelated Brazilian women at-risk for hereditary breast and ovarian cancer from the Oncogenetics Department of BCH. Those women were referred from the Oncogenetics Department of BCH for BRCA1, BRCA2 and TP53 genetic testing due to the presence of clinical criteria for HBOC, but no genetic alterations in these genes were found. For the purpose of the present study, were included only families fulfilling the following criteria: patients diagnosed with breast/ovarian cancer at an early age (<55 years), with at least two relatives with breast and/or ovarian cancer, two or more generations affected by cancer and absence of male BC.

Clinical information was obtained through detailed review of the patient´s clinical chart. For family history data, all pedigrees were revised.

Sequencing of *BRCA1*, *BRCA2* and *TP53*

Analysis of the presence of germline mutations *in BRCA1/BRCA2/TP53* genes was conducted at the Center of Molecular Diagnosis of BCH as part of routine care through NGS sequencing followed by rearrangement analysis through MLPA (Multiplex Ligation-dependent Probe Amplification Analysis), as described elsewhere by Fernandes *et al.* [49].

Tumor samples

For aCGH analysis, a representative section of FFPE tumor tissue from the breast or the ovarian tumor was stained by hematoxylin and eosin (H&E) and evaluated by a pathologist to verify tumor content (>70% tumor) and further microdissection.

DNA isolation and quality control

Following microdissection, DNA extraction steps were carried out using DNeasy Blood and Tissue kit (Qiagen), following the manufacturer's instructions. The quality and integrity of

the extracted DNA was assessed by multiplex PCR reaction using four primer pairs for the GAPDH gene (amplifying 100, 200, 300 and 400 bp, respectively), as described by Van Beers *et al.* [50]. The PCR reaction carried out contained (in a final volume of 30 µL) 1.5 mM MgCl2; 0.2 mM dNTP (Invitrogen); 0.133 µM of each primer; 1 U Taq DNA polymerase (Invitrogen) and 60 ng of tumor DNA. Reactions were performed in a Veriti thermocycler (Thermo Fisher Scientific) using the following amplification parameters: 94° C for 1 minute, 35 cycles of 94° C for 1 minute, 56° C for 1 minute, and 72° C for 3 minutes. Finally, a final extension at 72° C for 7 minutes. Amplification of DNA was verified by agarose gel electrophoresis.

Array comparative genomic hybridization

aCGH was performed on oligonucleotide-based SurePrint G3 Unrestricted CGH 8 × 60 K microarray slides, according the protocol provided by the manufacturer. In brief, 1 µg in final volume of 13 µL of normal female control DNA – reference DNA (DNA universal control-Promega Madison WI USA- Woman Reference: G152A) and patient's DNA were differentially labeled with Cy3 (cyanine 3-deoxyuridine triphosphate) and Cy5 (cyanine 5-deoxyuridina triphosphate), respectively, using Agilent SureTag Complete DNA Labeling Kit (Agilent Technologies). Labeled DNA was then cleaned with purification columns (Agilent Technologies) and hybridized on array at 65° C for 24 hours, according to manufacturer's recommendations. Microarrays were washed using Agilent Oligo aCGH Wash Buffers and scanning was performed using Agilent SureScan Microarray Scanner according to manufacturer's instructions (Agilent Technologies).

Data analysis

Data quantification of aCGH was performed with Feature Extraction software (Agilent Technologies) and the txt output files were imported into Nexus Copy Number v8.0 (BioDiscovery Inc) for visualization and downstream analysis. BioDiscovery's FASST2 Segmentation Algorithm, a Hidden Markov Model (HMM) based approach, was used to make copy number calls. The FASST2 algorithm, unlike other common HMM methods for copy number estimation, does not aim to estimate the copy number state at each probe but uses many states to cover more possibilities, such as mosaic events. These state values are then used to make calls based on a log-ratio threshold. The significance threshold for segmentation was set at 5.0E-6 also requiring a minimum of 3 probes per segment and a maximum probe

spacing of 1,000 kb between adjacent probes before breaking a segment. The log ratio thresholds for single copy gain (or amplification) and single copy loss (or deletion) were set at 0.2 and -0.23, respectively. The log ratio thresholds for two or more copy gain (or high copy gain) and homozygous loss (or high copy loss) were set at 1.14 and -1.1 respectively. A 3:1 sex chromosome gain threshold was set to 1.2 and a 4:1 sex chromosome gain threshold was set to 1.7. Male sex chromosome big loss threshold was set to -1.1. GISTIC (Genomic Identification of Significant Targets in Cancer) algorithm was used within Nexus 8.0 to identify regions that are significantly amplified or deleted across a set of samples. It was considered the default parameters of Q-bound ≤ 0.05 with False Discovery Rate (FDR) correction and G-score cut-off ≤ 1.0. The identification of genes and CNVs were also performed within Nexus 8.0, being CNVs filtered according to 1000 genomes project. It was calculated the frequency of the gained and lost remaining CNVs and further separated according to <1% (rare CNVs) and ≥1% (common CNVs). The peaks identified by GISTIC algorithm were associated to breast and ovarian cancer family history and clinical characteristics, i.e. clinical staging, age at diagnosis (≤30, 31–45 and ≥45 years), molecular subtype, histological subtype, presence of metastasis and recurrence. These analyses were done by Fisher's exact test (within SPSS v.21.0 software for Windows (Chicago, IL) considering the significance level of 5%.

Besides, the genomic regions found to be significant in GISTIC were considered for further analysis using the professional version of the compendium of cancer transcriptome profiles, Oncomine™ (Compendia Bioscience, Ann Arbor, MI). There were selected 13 breast and 5 ovarian cancer datasets (totalizing over 4000 samples). For each cancer type (breast or ovary), we selected the genes that presented gain or loss in our aCGH, and considered relevant those that presented gain in our aCGH and overexpression in Oncomine ($P < 0.01$), or those that presented loss in our aCGH and loss of expression in Oncomine ($P < 0.01$).

RESULTS

In the present study, we analyzed 31 Brazilian women at-risk for hereditary breast/ovarian cancer (27 with personal history of BC and 4 with ovarian tumors) without *BRCA1/BRCA2/TP53* germline mutations, by array-CGH. Clinicopathological characteristics and family history of the patients are specified in Table 1.

Briefly, the mean age at BC diagnosis was 42.9 years (SD = 7.9), ranging from 27–70 years. The majority of BC was invasive ductal carcinoma (77.8%), estrogen and progesterone

positive (69.2% and 73.1%, respectively) and HER2 negative (64.0%). Regarding molecular classification, the majority of patients presented luminal type tumors (21 patients, 80.8%), four patients (15.4%) had triple negative tumors and only one patient (3.8%) was diagnosed with a HER2 subtype tumor.

All four ovarian cancer patients developed serous adenocarcinoma subtype tumors. The average age at diagnosis was 47.7 years (SD = 18.0), ranging from 21–60 years.

A detailed cancer family history can be found in Table 1. All patients reported at least one case of BC in the family, diagnosed at early age (<55 years for breast cancer cases). In addition, two women at-risk for hereditary BC (samples: 960 and 1024) had a family history with bilateral BC. Among patients diagnosed with BC, the majority reported more than three cases of BC in the family history (16 cases, 59.3%). Meanwhile, all patients diagnosed with ovarian cancer, reported three or less BC cases in their families ($p = 0.043$). Moreover, 12 patients reported the presence of ovarian cancer in the family history.

Table 1: Clinicopathological characteristics and family history of the patients at-risk for hereditary cancer

Family	Cancer (age at diagnosis)	Histological type	Molecular subtype	Breast/Ovarian cancer cases in the family (sex and age at diagnosis, if known)
19	Breast (44)	IDC	ER: –; PR: +; HER2: –	Sister: Breast (F,46; F,46)
29	Ovarian (42), Breast (53)	DCIS	ER: +; PR: +	Paternal side of the family: Breast (F,29); Ovarian (F,60; F,?; F, ?), Uterus (F,57; F,?; F,?); Gastric (M,42; M,?; M,?; M,?)
65	Breast (35)	IDC	ER: –; PR: –; HER2: –	Maternal side of the family: Breast (F,31; F,34; F,47; F,39; F,39; F,46)
80	Breast (43)	DCIS	ER: +; PR: +; HER2: +	Maternal side of the family: Breast (F,44; f,44; F,55; F,57; F,60), Prostate (M,?)
85	Breast (51)	IDC	ER: +; PR:+; HER2: –	Maternal side of the family: Breast (F,43; F,45; F,48), Stomach (F;45; M,56); Leukemia (M,69)
179	Breast (47)	IDC	ER: +; PR: +	Paternal side of the family: Breast (F,37; F,49; F,61), Throat (M,?; M,?)
233	Breast (49)	IDC	ER: –; PR: –; HER2: –	Maternal side of the family: Breast (F,?; F,50; F,33; F,70; F,60; F,60; F,46), Colorectal (M,65), Gastric (M,62), Pancreas (M,62), Lung (M,52; M,66; M,?)
241	Breast (45)	IDC	ER: +; PR: +; HER2: –	Paternal side of the family: Breast (F,48; F,49)
275	Ovarian (60)	Serous adenocarcinoma	Not applicable	Paternal side of the family: Breast (F,32; F,35), Prostate (M,80)
289	Breast (48)	IDC	ER: –; PR: –; HER2: +	Maternal side of the family: Breast (F,50; F,65; F,65)
306	Melanoma (26), Breast (36)	IDC	ER: +; PR: +; HER2: –	Maternal side of the family: Breast (F,43; F,?)
320	Ovarian (53)	Serous adenocarcinoma	Not applicable	Maternal side of the family: Breast (F,52), Ovarian (F,71), Uterus (F,60), Thyroid (M,29), Lung (M,83)
426	Breast (38)	IDC	ER: +; PR: +; HER2: –	Paternal side of the family: Breast (F,20); Ovarian (F,28), Leukemia (M,78), Esophagus (M,?)
494	Breast (33)	IDC	ER: +; PR: +; HER2: +	Maternal side of the family: Breast (F,38; F,?), Ovarian (F,38)
558	Breast (37)	IDC	ER: –; PR: –; HER2: –	Maternal side of the family: Breast (F,52; F,?; F,?; F,?, F,?), Ovarian (F,42), Skin (F,?)
563	Breast (39)	IDC	ER: +; PR: –; HER2: –	Paternal side of the family: Breast (F,30; F,40; F,45, F,50; F,51), Lung, (M,?), Colorectal (F,64), Skin (M,72)
581	Breast (45)	DCIS	Not available	Paternal side of the family: Breast (F,49; F,46; F,54), Prostate (M,60; M,70), head and neck (M,83)
593	Breast (39)	IDC	ER: +; PR: +; HER2: –	Paternal side of the family: Breast (F,50; F,?; F,?; F,?), Ovarian (F,?; F,?), Gastric (M,?; M,?), Colorectal (M,?; M,?)
626	Breast (46)	DCIS	ER: +; PR: +	Maternal side of the family: Breast (F,74; F,80; F,57; F,45), Ovarian (F,45), Thyroid (F,40), Skin (M,80), Pancreas (M,?), Myeloma (M,60), Lips (M,?)
638	Breast (42)	ILC	ER: +; PR: +	Maternal side of the family: Breast (F,49; F,50; F,?), Gastric (F,55), Thyroid (F,36), Lips (F,55)

649	Breast (38)	IDC	ER: +; PR: +; HER2: –	Maternal side of the family: Breast (F,64), Ovarian (F,61), Thyroid (F,61)
695	Ovarian (21)	Serous adenocarcinoma	Not applicable	Paternal side of the family: Breast (F,42), Ovarian (F,68), Colorectal (F,40; M,40), Gastric (F,50; F,70)
960	Bilateral Breast (59,70)	IDC	ER: +; PR: +; HER2: –	Maternal side of the family: Breast (F,34; F,59), Uterus (F,45), Lung (M,77; M,?)
974	Breast (46)	IDC	ER: +; PR: +; HER2: –	Maternal side of the family: Breast (F,55; F,45; F,60; F,60; F,55; F,45; F,60), Prostate (M,70; M,80)
981	Breast (37)	IDC	ER: +; PR: +; HER2: –	Maternal side of the family: Breast (F,32; F,70; F,60), Melanoma (F,30; F36), Leukemia (F,5), Bile ducts (M,49; F,55)
1014	Breast (42)	DCIS	ER: +; PR: +	Maternal side of the family: Breast (F,53; F,?); Melanoma (F,75), Lymphoma (M,19), Liver (F,?), Brain (F,?)
1024	Breast (48)	IDC	ER: –; PR: –; HER2: –	Paternal side of the family: Breast (F,70; F,72; F,44; F;44; F,49), Ovarian (F,56), Colorectal (M,20), Melanoma (M, ?), Prostate (M,50), Gastric (F,70; F,72; F,41)
1055	Ovarian (57)	Serous adenocarcinoma	Not applicable	Maternal side of the family: Breast (F,49; F,50), Pancreas (F,50), Lung (M,?)
1095	Breast (43)	IDC	ER: –; PR: –; HER2: +	Paternal side of the family: Breast (F,27; F,42), Uterus (F,98), Throat (M,72)
1151	Breast (38)	IDC	ER: +; PR: +; HER2: –	Maternal side of the family: Breast (F,35; F,60)
1264	Breast (27)	IDC	ER: –; PR: +; HER2: –	Maternal side of the family: Breast (F,50), Pancreas (M,75); Intestine (M,81)

DCIS: ducatal carcinoma *in situ*; ER: estrogen receptor; F: female; IDC: Invasive ductal carcinoma; M: male; PR: progesterone receptor.

The molecular analysis revealed gained and lost regions across all chromosomes for both breast and ovarian tumors (Figure 1). We found 20 gained regions and 31 lost in BRCAX tumors. In addition, some variations, although not statistically significant, were found only in patients diagnosed with breast cancer, such as: gains of 7p22.1, 12p13.1, 14q13.3-q21.1, 17q11.2, 17q12 and 17q21.32-q21.33, and losses of 2p25.3, 6q25.3-q26 and 10q26.3. Moreover, the gain of Xq26 and loss of Xp22.32– 22.31 were more frequent in ovarian cancer (100%), compared with breast cases (26% and 59%, respectively) ($p = 0.01$ for both regions). Loss of 22q13.31–13.32 was detected more often in ovarian than in breast cancer cases ($p = 0.043$). In addition, a significant number of copy number alterations involving chromosome 8 was observed.

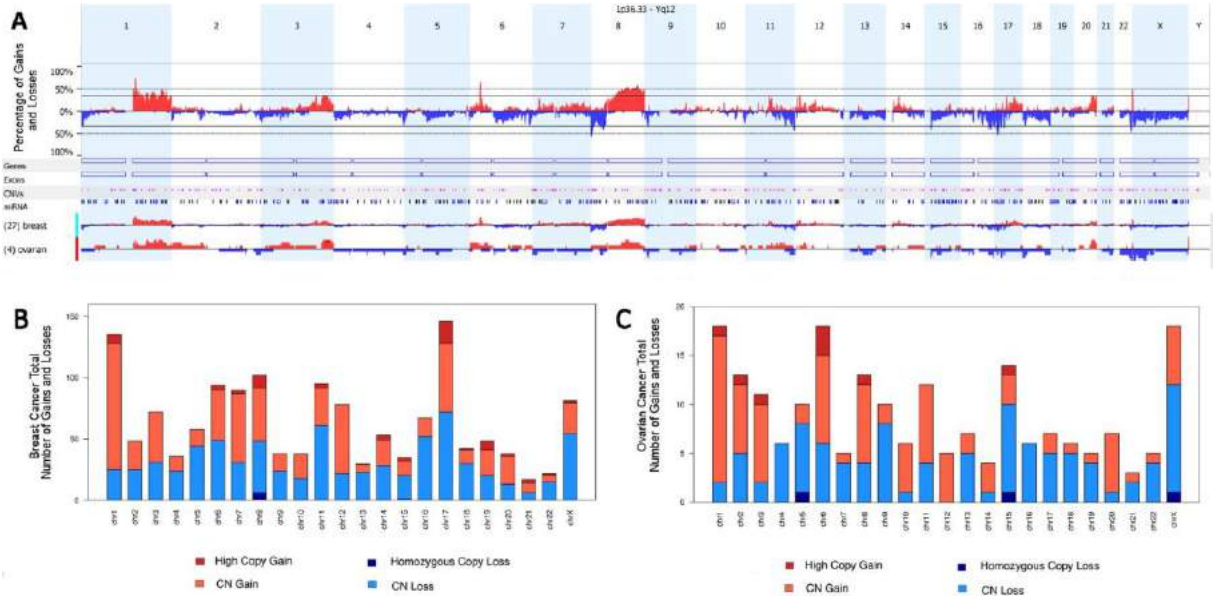


Figure 1: Overview of gained and lost regions across all chromosomes. **(A)** Overall and specific breast and ovarian copy number aberration frequencies. Regions presenting copy gains are shown in red and with copy loss in blue. **(B)** Overview of gained and lost regions across all chromosomes in breast tumors. **(C)** Overview of gained and lost regions across all chromosomes in ovarian tumors.

When family history was taken into consideration for copy number variation analyses, we observed that loss of 22q13.31–13.32 region was significantly associated with the presence of ovarian family history (p = 0.03). This region includes *MIR3201*, *LOC284933*, *FAM19A5*, *MIR4535*, *LINC01310* genes. Other significant association found included gains in the 6p22.1 region (including 13 histone family genes) in 100% of metastatic cases (p = 0.03). Finally, we found loss of 6q25.1 in 71% of patients with metastasis (p = 0.01). This region includes *RAET1E*, *RAET1E-AS1*, *RAET1G*, *ULBP2*, *ULBP1*, *RAET1K*, *RAET1L*, *ULBP3*, *PPP1R14C*, *IYD*, *PLEKHG1*, *MTHFD1L* genes (Supplementary Table 1).

In addition, when comparing our findings with those of the literature of BRCAX tumors, we observed that our results corroborate some findings reported by Didraga *et al*. (2011), Alvarez *et al*. (2016) and Mangia *et al*. (2008), showing 50%, 21% and 12% of common regions, respectively (Figure 2).

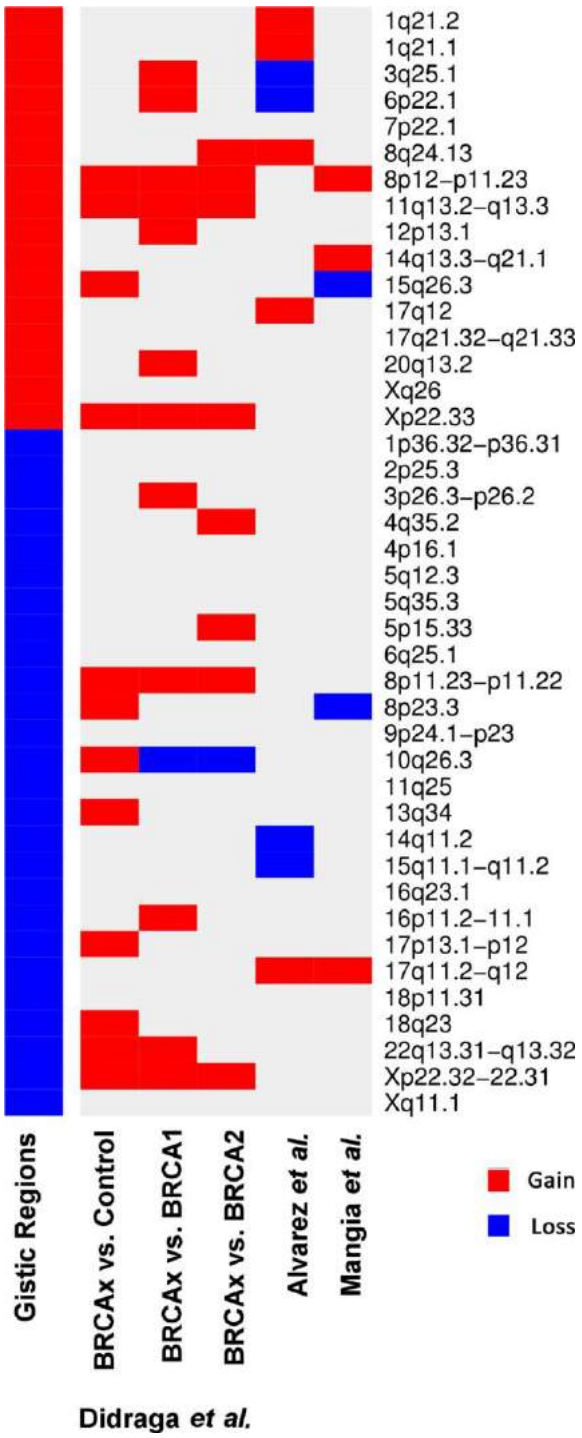


Figure 2: Heatmap representing the gains (in red) and losses (in blue) through aCGH found by GISTIC algorighm in common with previous studies by Didraga *et al.* (2011), Alvarez *et al.* (2016) and Mangia *et al.* (2008).

Finally, we found that 22 genes present in gained regions also present overexpression in the Oncomine database, whereas 21 genes present in lost regions show loss of expression in the same database ($p < 0.01$, Table 2).

Table 2: Genes in gained regions that presented *in silico* overexpression and genes in lost regions that presented *in silico* loss of expression.

Event[1]	Cytoband	Genes
Gain/Overexp	1q21.1-q21.2	*PEX11B, PDE4DIP, ECM1, TARS2, RPRD2*
Gain/Overexp	6p22.1	*HIST1H3H, HIST1H4J, HIST1H4K*
Gain/Overexp	7p21.1	*HDAC9*
Gain/Overexp	8p11.23-p11.22	*TM2D2, LETM2, RNF5P1, FGFR1*
Gain/Overexp	8q24.13	*NSMCE2, KIAA0196*
Gain/Overexp	11q13.3	*CTTN*
Gain/Overexp	17q12	*ERBB2, STARD3, GRB7*
Gain/Overexp	17q21.32-q21.33	*PHB, ABI3*
Gain/Overexp	Xp22.33	*CRLF2*
Loss/LOExp	1p36.32	*TPRG1L, AJAP1*
Loss/LOExp	2p25.3	*FAM150B, TMEM18, TPO*
Loss/LOExp	3p26.3-p26.2	*CRBN, CNTN4*
Loss/LOExp	5q35.3	*ADAMTS2, ZNF879, COL23A1*
Loss/LOExp	8p23.3	*ERICH1, RPL23AP53, OR4F21, ZNF596*
Loss/LOExp	11q25	*JAM3, LOC283177, THYN1*
Loss/LOExp	14q11.2	*OR4K5*
Loss/LOExp	16q23.1	*CNTNAP4, SYCE1L*
Loss/LOExp	Xp22.32-p22.31	*NLGN4X*

[1]Events represent Gain or Loss in our samples with concurrent Overexpression (Overexp) or Loss of Expression (LOExp) on Oncomine samples.

DISCUSSION

In the present study, a BRCAX tumor characterization of FFPE samples has been performed by array comparative genomic hybridization. Among the altered loci, we can highlight the identification of several alterations in chromosome 8, including losses on 8p11.23-p11.22 and gains on 8p12-p11.23 and 8q24.13, in concordance with previous studies of BRCAX tumors [14, 25, 26]. Besides, the chromosomal region 8p12-p11 has been reported to be amplified in 10–23% of BC cases [27–29], and some studies have shown that amplification on this region is associated with poor clinical outcome [27, 30]. We found by *in silico* analysis that 4 genes present in this region (including *FGFR1* and *NSMCE2*) are overexpressed.

The *FGFR1* gene encodes a transmembrane protein that interacts with fibroblast growth factors and directly influence mitogenesis and cell differentiation. In fact, there are several studies showing different treatment outcomes of breast cancer women depending on

the FGFR1 status [31–33]. Similarly, *NSMCE2* plays an important role in cell cycle, since its depletion in MCF-7 breast cancer cells affected cell cycle and G1-S transition [34]. Moreover, the overexpression of cortactin (*CTTN*), present in 11q13.3, was linked to *CCND1* amplification in premenopausal breast cancer [35], although it failed to demonstrate a strong prognostic value in patients with breast cancer [36]. Conversely, its upregulation promoted colon cancer progression through ERK pathway [37]. Therefore, other studies have shown that amplification on chromosomal region 8p12-p11 in combination with amplification on 11q13 have more impact on patient outcome than amplification on only one of the two loci [27, 38].

In addition to gains and losses on chromosome 8 and alterations on chromosome 11, alterations in chromosome X seem to be characteristic of BRCAX tumors. In our study, a great number of samples showed gains on regions 11q13.2-q13.3 and Xp22.33, which were also identified by Didraga and collaborators [25]. Although it is not extensively studied in breast cancer, the overexpression of *CRLF2*, present in Xp22.33, has been demonstrated to be a marker of poor outcome of pediatric and adult B-precursor acute lymphoblastic leucemia (ALL) (as reviewed in [39]).

Study performed by Gronwald *et al*. [19] compared BRCAX with sporadic breast cancers and identified several altered regions (114 gains and 36 losses) in 18 patients. Their findings showed concordances with our results, presenting more often gains in 1q, 6p, 17q and frequent loss of 8p. Beside the well-known effects of *ERBB2* amplification in breast cancer development, the overexpression of *STARD3* (located in the same locus) seems to be important, since it may contribute to increased proliferation, migration and invasion of breast cancer cells (as reviewed in [40]). Finally, considering our findings of altered regions found in BRCAX associated with metastasis (gain of 6p22.1 and loss of 6q25.1), gain of 6p was previously associated with BRCAX, and loss of 6q with BRCA1 tumors [41]. In fact, there are several members of RAET and ULBP family present in this locus. These members are ligands of C-type lectin-like receptor NKG2D, present in NK and T cells subsets, highly involved in tumor immunosurveillance [42]. Therefore, the loss of this region may have led to lower expression of these ligands, leading to less immunogenicity of the tumor cells. In fact, there are reports on colorectal cancer that have demonstrated this same pattern, and several authors discuss the potential therapeutic utility of NKG2D ligands in the treatment of this disease [42–44]. Therefore, these alterations on chromosome 6 seem to be highly associated with breast cancer tumors and may be of interest for further studies.

We also found that loss of 22q13.31–13.32 was significantly associated with presence of ovarian tumors (in the proband or in the family). The loss of heterozygosis (LOH) of chromosome 22q has been reported in a variety of cancers, including ovarian cancers, where the LOH rates reached 70% of cases [45, 46]. Study published by Zweemer *el al.* (2001) reported a significant loss of 22q, identified through aCGH in familial ovarian tumors [47]. Interestingly, *MIR3201* was significantly downregulated in recurrent epithelial ovarian cancer (EOC), when compared to primary EOC [48]. To the best of our knowledge, there are no studies pointing to the functional relevance of *MIR3201* in ovarian cancer, however, further studies may be performed to evaluate its possible role as a biomarker of EOC recurrence.

In summary, our findings support previous data of BRCAX related alterations and point to new regions potentially associated with personal and family history of ovarian cancer. In the present study, we could identify by aCGH analysis a potentially BRCAX-associated ovarian region on chromosome 22. Given our limited sample size, further work should be performed in order to validate our findings, to identify the driver genes associated with the BRCAX tumor development, as well as to uncover the role of those altered regions in cancer formation and progression.

Abbreviations

aCGH: Array comparative genomic hybridization; BC: breast cancer; BCH: Barretos Cancer Hospital; ER: estrogen receptor; HBOC: Hereditary Breast and Ovarian Cancer Predisposition Syndrome; HER2: Human Epidermal growth factor Receptor 2; PR: progesterone receptor.

ACKNOWLEDGMENTS

The authors thank the patients, the Pathology Department and the Center of Molecular Diagnosis of Barretos Cancer Hospital.

CONFLICTS OF INTEREST

The authors declare that they have no competing interests.

FUNDING

This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP - 2013/24633-2). EIP is recipient of a National Council of Technological and Scientific Development (CNPq) scholarship.

Supplementary Table 1 - Complete list of candidate genes and CNVs by GISTIC

Gistic regions	Event	Number of samples	Genes	CNVs < 1%	CNVs > 1%
chr1:148,555,037-148,930,370	Gain	21	PRPF3,RPRD2,MIR6878,TARS2,ECM1,FALEC,MIR4257,ADAMTSL4,ADAMTSL4-AS1,MCL1,ENSA,GOLPH3L	-	-
chr1:143,639,135-144,278,309	Gain	16	NBPF20,PDE4DIP,NBPF12,SEC22B,NOTCH2NL,NBPF25P,LOC101928979,HFE2,TXNIP,POLR3GL,ANKRD34A,LIX1L,RBM8A,GNRHR2,PEX11B,ITGA10,ANKRD35,NBPF10	esv3380599,esv3414198,esv3360116,esv3368223,esv3352814,esv3387252,esv3361392,esv3349227,dgv191e59	dgv199e59,dgv200e59,esv3355910,esv3309997,esv3304560,esv3405247,esv3413349,dgv201e59
chr3:151,524,863-152,116,104	Gain	11	TSC22D2,SERP1,EIF2A,SELT,ERICH6,ERICH6-AS1,LOC101928105,SIAH2	esv3342752	-
chr6:27,864,081-27,957,537	Gain	20	HIST1H2BL,HIST1H2AI,HIST1H3H,HIST1H2AJ,HIST1H2BM,HIST1H4J,HIST1H4K,HIST1H2AK,HIST1H2BN,HIST1H2AL,HIST1H1B,HIST1H3I,HIST1H4L	-	-
chr7:17,587,293-18,535,820	Gain	9	SNX13,PRPS1L1,MIR1302-6,HDAC9	esv3394688,esv3362182,esv3306132,esv3380997	dgv3749e59,esv3374815,esv3306306,esv3354676,esv3389040,esv3306932,esv3308523,esv3304893,esv3427043,esv3441757,esv3437954
chr8:126,015,907-126,691,570	Gain	18	LINC00964,ZNF572,SQLE,KIAA0196,NSMCE2,TRIB1	-	esv3342637,esv3445339
chr8:38,233,025-39,318,619	Gain	5	DDHD2,PLPP5,WHSC1L1,LETM2,FGFR1,C8orf86,RNF5P1,TACC1,PLEKHA2,HTRA4,TM2D2,ADAM9,ADAM32,ADAM5	esv3433731,esv3449093	dgv4129e59,esv3359463
chr11:69,849,987-70,077,870	Gain	9	PPFIA1,CTTN,SHANK2	-	esv3381709

chr12:12,630,063-12,788,197	Gain	11	CREBL2,GPR19,CDKN1B,APOLD1	-	-
chr14:36,796,282-37,159,041	Gain	10	MIPOL1,FOXA1,TTC6	-	esv3398904
chr15:96,667,275-97,452,797	Gain	2	FAM169B,IRAIN,MIR4714,IGF1R,PGPEP1L,LUNAR1	esv3435433,esv3404889,esv3361098	esv3408735
chr17:35,060,044-35,156,109	Gain	10	STARD3,TCAP,PNMT,PGAP3,ERBB2,MIR4728,MIEN1,GRB7	-	-
chr17:44,648,530-44,976,930	Gain	6	ABI3,PHOSPHO1,FLJ40194,MIR6129,ZNF652,LOC102724596,PHB,LOC101927207,NGFR,MIR6165,LOC100288866	-	esv3326018, esv3307703
chr20:50,878,094-51,142,697	Gain	10	TSHZ2	-	-
chrX:154,638,096-154,913,754	Gain	11	SPRY3,VAMP7,IL9R,DDX11L16	-	-
chrX:0-1,514,088	Gain	15	PLCXD1,GTPBP6,PPP2R3B,SHOX,CRLF2,CSF2RA,MIR3690,IL3RA,SLC25A6,LINC00106,ASMTL-AS1,ASMTL	dgv4550e59,dgv4551e59,dgv4552e59,dgv4553e59,dgv4554e59,dgv4555e59,dgv4556e59,dgv4557e59,dgv4558e59,dgv4559e59,dgv4560e59,dgv4561e59,dgv4562e59	esv3334914,esv3414524,esv3450882,esv3382952,esv3442681,esv3426006,esv3327828,esv3447355,esv3447400,esv3338982,esv3439707,esv3310495,esv3336590,esv3389029,esv3416240,esv3334895,esv3362005,esv3405532,esv3448400,esv3326708
chr1:2,511,532-6,068,705	Loss	11	FAM213B,MMEL1,TTC34,ACTRT2,LINC00982,MIR4251,PRDM16,ARHGEF16,MEGF6,MIR551A,TPRG1L,WRAP73,TP73,TP73-AS1,CCDC27,SMIM1,LRRC47,CEP104,DFFB,C1orf174,LINC01134,LINC01346,LOC284661,AJAP1,MIR4417,MIR4689,NPHP4,KCNAB2	-	-

chr2:0-3,901,873	Loss	7	FAM110C,SH3YL1,ACP1,FAM150B,TMEM18,LINC01115,LOC101060385,SNTG2,TPO,PXDN,MYT1L,MYT1L-AS1,LINC01250,TSSC1,TRAPPC12,ADI1,RNASEH1,RNASEH1-AS1,RPS7,COLEC11,ALLC,DCDC2C	-	-
chr3:2,363,239-4,305,022	Loss	8	CNTN4,CNTN4-AS1,IL5RA,TRNT1,CRBN,LRRN1	-	-
chr4:190,034,368-191,273,063	Loss	10	LINC01262,LINC01596,LOC283788,FRG1,FRG1CP,FRG2,DBET	-	-
chr4:7,135,530-7,702,329	Loss	7	FLJ36777,MIR4798,PSAPL1,MIR4274,SORCS2	-	-
chr5:65,956,884-66,166,332	Loss	7	MAST4	-	-
chr5:177,726,909-178,886,217	Loss	7	COL23A1,CLK4,ZNF354A,AACSP1,ZNF354B,ZFP2,ZNF454,GRM6,ZNF879,ZNF354C,ADAMTS2	-	-
chr5:2,075,794-3,822,399	Loss	8	LOC100506858,IRX2,C5orf38,LINC01377,LINC01019,LINC01017,IRX1	-	-
chr6:150,238,707-151,397,838	Loss	9	RAET1E,RAET1E-AS1,RAET1G,ULBP2,ULBP1,RAET1K,RAET1L,ULBP3,PPP1R14C,IYD,PLEKHG1,MTHFD1L	-	-
chr8:39,359,817-39,533,168	Loss	10	ADAM5,ADAM3A	-	-
chr8:0-1,452,246	Loss	18	OR4F21,RPL23AP53,ZNF596,FAM87A,FBXO25,TDRP,ERICH1,LOC401442,ERICH1-AS1,LOC286083,DLGAP2	-	-
chr9:8,381,903-10,135,588	Loss	11	PTPRD-AS1,PTPRD	-	-

chr10:131,710,606-134,068,182	Loss	9	LINC00959,CTAGE7P,GLRX3,MIR378C,TCERG1L-AS1,TCERG1L,LINC01164,PPP2R2D,BNIP3,JAKMIP3,DPYSL4,STK32C,LRRC27,PWWP2B	-	-
chr11:132,714,177-134,452,384	Loss	14	OPCML,LOC646522,SPATA19,MIR4697,MIR4697HG,IGSF9B,LOC100128239,JAM3,NCAPD3,VPS26B,THYN1,ACAD8,GLB1L3,GLB1L2,B3GAT1,LOC283177	-	-
chr13:110,713,667-111,877,154	Loss	7	ARHGEF7,TEX29,LINC00354,LINC00403,SOX1	-	-
chr14:19,323,579-19,564,886	Loss	8	OR4N2,OR4K2,OR4K5,OR4K1,OR4K15,OR4K14	-	-
chr15:18,362,555-19,906,749	Loss	13	CHEK2P2,HERC2P3,GOLGA6L6,GOLGA8CP,NBEAP1,LOC646214,CXADRP2,MIR3118-2,MIR3118-3,MIR3118-4,POTEB2,POTEB,POTEB3,NF1P2,MIR5701-1,MIR5701-2,MIR5701-3,LINC01193,LOC727924,LOC101927079,OR4M2,OR4N4	-	-
chr16:75,113,428-76,314,342	Loss	13	CNTNAP4,LOC101928203,MIR4719,MON1B,SYCE1L,ADAMTS18,NUDT7	-	-
chr16:33,454,467-34,903,000	Loss	8	ENPP7P13,LINC00273,UBE2MP1,LINC01566,FRG2DP,TP53TG3HP,FLJ26245	-	-
chr17:10,832,111-11,804,121	Loss	17	SHISA6,DNAH9	-	-
chr17:28,641,855-30,271,750	Loss	12	AA06,ASIC2,LOC101927239,CCL2,CCL7,CCL11,CCL8,CCL13,CCL1,C17orf102,TMEM132E	-	-

chr18:5,961,056-7,125,051	Loss	10	L3MBTL4,L3MBTL4-AS1,MIR4317,LINC01387,ARHGAP28,LINC00668,LOC101927188,LAMA1	-	-
chr18:73,158,585-74,939,833	Loss	11	LINC01029,SALL3,ATP9B	-	-
chr22:46,795,328-48,007,549	Loss	15	MIR3201,LOC284933,FAM19A5,MIR4535,LINC01310	-	-
chrX:5,127,672-8,738,437	Loss	11	NLGN4X,MIR4770,VCX3A,PUDP,MIR4767,STS,VCX,PNPLA4,MIR651,VCX2,VCX3B,ANOS1,FAM9A	-	-
chrX:62,893,204-64,474,249	Loss	10	ARHGEF9,MIR1468,AMER1,ASB12,MTMR8,ZC4H2	-	-

REFERENCES

1. WHO. World Health Organization - Breast Cancer. 2016. available from: http://www.who.int/cancer/detection/ breastcancer/en/.

2. Lynch HT, Snyder C, Lynch J. Hereditary breast cancer: practical pursuit for clinical translation. Ann Surg Oncol. 2012; 19:1723–31. https://doi.org/10.1245/ s10434-012-2256-z.

3. Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, Cochran C, Bennett LM, Ding W, Bell R, Rosenthal J, Hussey C, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science. 1994; 266:66–71.

4. Wooster R, Neuhausen SL, Mangion J, Quirk Y, Ford D, Collins N, Nguyen K, Seal S, Tran T, Averill D. Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. Science. 1994; 265:2088–90.

5. Gronwald J, Byrski T, Huzarski T, Oszurek O, Janicka A, Szymanska-Pasternak J, Gorski B, Menkiszak J, RzepkaGorska I, Lubinski J. Hereditary breast and ovarian cancer. Hered Cancer Clin Pract. 2008; 6:88–98. https://doi. org/10.1186/1897-4287-6-2-88.

6. Petrucelli N, Daly MB, Pal T. BRCA1- and BRCA2-Associated Hereditary Breast and Ovarian Cancer. In: Pagon RA, Adam MP, Ardinger HH, Wallace SE, Amemiya A, Bean LJH, Bird TD, Ledbetter N, Mefford HC, Smith RJH and Stephens K, eds. GeneReviews(R). (Seattle (WA). 1993. Available from: https:// www.ncbi.nlm.nih.gov/books/NBK1247/.

7. Friedenson B. BRCA1 and BRCA2 pathways and the risk of cancers other than breast or ovarian. MedGenMed. 2005; 7:60.

8. Couch FJ, Nathanson KL, Offit K. Two decades after BRCA: setting paradigms in personalized cancer care and prevention. Science. 2014; 343:1466–70. https://doi. org/10.1126/science.1251827.

9. Mavaddat N, Antoniou AC, Easton DF, Garcia-Closas M. Genetic susceptibility to breast cancer. Mol Oncol. 2010; 4:174–91. https://doi.org/10.1016/j.molonc.2010.04.011.

10. Tanic M, Andres E, Rodriguez-Pinilla SM, MarquezRodas I, Cebollero-Presmanes M, Fernandez V, Osorio A, Benitez J, Martinez-Delgado B. MicroRNA-based molecular classification of non-BRCA1/2 hereditary breast tumours. Br J Cancer. 2013; 109:2724–34. https://doi. org/10.1038/bjc.2013.612.

11. Honrado E, Osorio A, Milne RL, Paz MF, Melchor L, Cascon A, Urioste M, Cazorla A, Diez O, Lerma E, Esteller M, Palacios J, Benitez J. Immunohistochemical classification of non-BRCA1/2 tumors identifies different groups that demonstrate the heterogeneity of BRCAXfamilies. Mod Pathol. 2007; 20:1298–306. https://doi. org/10.1038/modpathol.3800969.

12. Hopper JL. More breast cancer genes? Breast Cancer Res. 2001; 3:154–7.

13. Oldenburg RA, Meijers-Heijboer H, Cornelisse CJ, Devilee P. Genetic susceptibility for breast cancer: how many more genes to be found? Crit Rev Oncol Hematol. 2007; 63:125–49. https://doi.org/10.1016/j. critrevonc.2006.12.004.

14. Mangia A, Chiarappa P, Tommasi S, Chiriatti A, Petroni S, Schittulli F, Paradiso A. Genetic heterogeneity by comparative genomic hybridization in BRCAx breast cancers. Cancer Genet Cytogenet. 2008; 182:75–83. https:// doi.org/10.1016/j.cancergencyto.2008.01.002.

15. Lacroix M, Leclercq G. The “portrait” of hereditary breast cancer. Breast Cancer Res Treat. 2005; 89:297–304. https:// doi.org/10.1007/s10549-004-2172-4.

16. Rahman N, Teare MD, Seal S, Renard H, Mangion J, Cour C, Thompson D, Shugart Y, Eccles D, Devilee P, Meijers H, Nathanson KL, Neuhausen SL, et al. Absence of evidence for a familial breast cancer susceptibility gene at chromosome 8p12-p22. Oncogene. 2000; 19:4170–3. https://doi.org/10.1038/sj.onc.1203735.

17. Imbert A, Chaffanet M, Essioux L, Noguchi T, Adelaide J, Kerangueven F, Le Paslier D, Bonaiti-Pellie C, Sobol H, Birnbaum D, Pebusque MJ. Integrated map of the chromosome 8p12-

p21 region, a region involved in human cancers and Werner syndrome. Genomics. 1996; 32:29–38.

18. Nathanson KL, Wooster R, Weber BL. Breast cancer genetics: what we know and what we need. Nat Med. 2001; 7:552–6. https://doi.org/10.1038/87876.

19. Gronwald J, Jauch A, Cybulski C, Schoell B, BohmSteuer B, Lener M, Grabowska E, Gorski B, Jakubowska A, Domagala W, Chosia M, Scott RJ, Lubinski J. Comparison of genomic abnormalities between BRCAX and sporadic breast cancers studied by comparative genomic hybridization. Int J Cancer. 2005; 114:230–6. https://doi. org/10.1002/ijc.20723.

20. Jonsson G, Naylor TL, Vallon-Christersson J, Staaf J, Huang J, Ward MR, Greshock JD, Luts L, Olsson H, Rahman N, Stratton M, Ringner M, Borg A, et al. Distinct genomic profiles in hereditary breast tumors identified by array-based comparative genomic hybridization. Cancer Res. 2005; 65:7612–21. https://doi.org/10.1158/0008-5472. CAN-05-0570.

21. Naylor TL, Greshock J, Wang Y, Colligon T, Yu QC, Clemmer V, Zaks TZ, Weber BL. High resolution genomic analysis of sporadic breast cancer using array-based comparative genomic hybridization. Breast Cancer Res. 2005; 7:R1186-98. https://doi.org/10.1186/bcr1356.

22. Nessling M, Richter K, Schwaenen C, Roerig P, Wrobel G, Wessendorf S, Fritz B, Bentz M, Sinn HP, Radlwimmer B, Lichter P. Candidate genes in breast cancer revealed by microarray-based comparative genomic hybridization of archived tissue. Cancer Res. 2005; 65:439–47.

23. Fridlyand J, Snijders AM, Ylstra B, Li H, Olshen A, Segraves R, Dairkee S, Tokuyasu T, Ljung BM, Jain AN, McLennan J, Ziegler J, Chin K, et al. Breast tumor copy number aberration phenotypes and genomic instability. BMC Cancer. 2006; 6:96. https://doi.org/10.1186/1471-2407-6-96.

24. van Beers EH, van Welsem T, Wessels LF, Li Y, Oldenburg RA, Devilee P, Cornelisse CJ, Verhoef S, Hogervorst FB, van't Veer LJ, Nederlof PM. Comparative genomic hybridization

profiles in human BRCA1 and BRCA2 breast tumors highlight differential sets of genomic aberrations. Cancer Res. 2005; 65:822–7.

25. Didraga MA, van Beers EH, Joosse SA, Brandwijk KI, Oldenburg RA, Wessels LF, Hogervorst FB, Ligtenberg MJ, Hoogerbrugge N, Verhoef S, Devilee P, Nederlof PM. A non-BRCA1/2 hereditary breast cancer sub-group defined by aCGH profiling of genetically related patients. Breast Cancer Res Treat. 2011; 130:425–36. https://doi. org/10.1007/s10549-011-1357-x.

26. Alvarez C, Aravena A, Tapia T, Rozenblum E, Solis L, Corvalan A, Camus M, Alvarez M, Munroe D, Maass A, Carvallo P. Different Array CGH profiles within hereditary breast cancer tumors associated to BRCA1 expression and overall survival. BMC Cancer. 2016; 16:219. https://doi. org/10.1186/s12885-016-2261-x.

27. Letessier A, Sircoulomb F, Ginestier C, Cervera N, Monville F, Gelsi-Boyer V, Esterni B, Geneix J, Finetti P, Zemmour C, Viens P, Charafe-Jauffret E, Jacquemier J, et al. Frequency, prognostic impact, and subtype association of 8p12, 8q24, 11q13, 12p13, 17q12, and 20q13 amplifications in breast cancers. BMC Cancer. 2006; 6:245. https://doi.org/10.1186/1471-2407-6-245.

28. Yang ZQ, Streicher KL, Ray ME, Abrams J, Ethier SP. Multiple interacting oncogenes on the 8p11-p12 amplicon in human breast cancer. Cancer Res. 2006; 66:11632–43. https://doi.org/10.1158/0008-5472.CAN-06-2946.

29. Bernard-Pierrot I, Gruel N, Stransky N, Vincent-Salomon A, Reyal F, Raynal V, Vallot C, Pierron G, Radvanyi F, Delattre O. Characterization of the recurrent 8p11-12 amplicon identifies PPAPDC1B, a phosphatase protein, as a new therapeutic target in breast cancer. Cancer Res. 2008; 68:7165–75. https://doi.org/10.1158/0008-5472. CAN-08-1360.

30. Gelsi-Boyer V, Orsetti B, Cervera N, Finetti P, Sircoulomb F, Rouge C, Lasorsa L, Letessier A, Ginestier C, Monville F, Esteyries S, Adelaide J, Esterni B, et al. Comprehensive profiling of 8p11-12 amplification in breast cancer. Mol Cancer Res. 2005; 3:655–67. https://doi. org/10.1158/1541-7786.MCR-05-0128.

31. Chen H, Singh RR, Lu X, Huo L, Yao H, Aldape K, Abraham R, Virani S, Mehrotra M, Mishra BM, Bousamra A, Albarracin C, Wu Y, et al. Genome-wide copy number aberrations and HER2 and FGFR1 alterations in primary breast cancer by molecular inversion probe microarray. Oncotarget. 2017; 8:10845–57. https://doi. org/10.18632/oncotarget.14802.

32. Luo J, Liu S, Leung S, Gru AA, Tao Y, Hoog J, Ho J, Davies SR, Allred DC, Salavaggione AL, Snider J, Mardis ER, Nielsen TO, et al. An mRNA Gene Expression-Based Signature to Identify FGFR1-Amplified Estrogen ReceptorPositive Breast Tumors. J Mol Diagn. 2017; 19:147–61. https://doi.org/10.1016/j.jmoldx.2016.09.007.

33. Shi YJ, Tsang JY, Ni YB, Chan SK, Chan KF, Tse GM. FGFR1 is an adverse outcome indicator for luminal A breast cancers. Oncotarget. 2016; 7:5063–73. https://doi. org/10.18632/oncotarget.6563.

34. Ni HJ, Chang YN, Kao PH, Chai SP, Hsieh YH, Wang DH, Fong JC. Depletion of SUMO ligase hMMS21 impairs G1 to S transition in MCF-7 breast cancer cells. Biochim Biophys Acta. 2012; 1820:1893–900. https://doi. org/10.1016/j.bbagen.2012.08.002.

35. Lundgren K, Holm K, Nordenskjold B, Borg A, Landberg G. Gene products of chromosome 11q and their association with CCND1 gene amplification and tamoxifen resistance in premenopausal breast cancer. Breast Cancer Res. 2008; 10:R81. https://doi.org/10.1186/bcr2150.

36. Sheen-Chen SM, Huang CY, Liu YY, Huang CC, Tang RP. Cortactin in breast cancer: analysis with tissue microarray. Anticancer Res. 2011; 31:293–7.

37. Ni QF, Yu JW, Qian F, Sun NZ, Xiao JJ, Zhu JW. Cortactin promotes colon cancer progression by regulating ERK pathway. Int J Oncol. 2015; 47:1034–42. https://doi. org/10.3892/ijo.2015.3072.

38. Karlsson E, Waltersson MA, Bostner J, Perez-Tenorio G, Olsson B, Hallbeck AL, Stal O. High-resolution genomic analysis of the 11q13 amplicon in breast cancers identifies synergy with 8p12 amplification, involving the mTOR targets S6K2 and 4EBP1. Genes Chromosomes Cancer. 2011; 50:775–87. https://doi.org/10.1002/gcc.20900.

39. Tasian SK, Loh ML. Understanding the biology of CRLF2overexpressing acute lymphoblastic leukemia. Crit Rev Oncog. 2011; 16:13–24.

40. Jacot W, Fiche M, Zaman K, Wolfer A, Lamy PJ. The HER2 amplicon in breast cancer: Topoisomerase IIA and beyond. Biochim Biophys Acta. 2013; 1836:146–57. https://doi. org/10.1016/j.bbcan.2013.04.004.

41. Kamieniak MM, Munoz-Repeto I, Rico D, Osorio A, Urioste M, Garcia-Donas J, Hernando S, Robles-Diaz L, Ramon YCT, Cazorla A, Saez R, Garcia-Bueno JM, Domingo S, et al. DNA copy number profiling reveals extensive genomic loss in hereditary BRCA1 and BRCA2 ovarian carcinomas. Br J Cancer. 2013; 108:1732–42. https://doi.org/10.1038/bjc.2013.141.

42. Rothe A, Jachimowicz RD, Borchmann S, Madlener M, Kessler J, Reiners KS, Sauer M, Hansen HP, Ullrich RT, Chatterjee S, Borchmann P, Yazaki P, Koslowsky TC, et al. The bispecific immunoligand ULBP2-aCEA redirects natural killer cells to tumor cells and reveals potent antitumor activity against colon carcinoma. Int J Cancer. 2014; 134:2829–40. https://doi.org/10.1002/ijc.28609.

43. Ramutton T, Buccheri S, Dieli F, Todaro M, Stassi G, Meraviglia S. γδ T cells as a potential tool in colon cancer immunotherapy. Immunotherapy. 2014; 6:989–99. https:// doi.org/10.2217/imt.14.59.

44. Xia Y, Chen B, Shao X, Xiao W, Qian L, Ding Y, Ji M, Gong W. Treatment with a fusion protein of the extracellular domains of NKG2D to IL-15 retards colon cancer growth in mice. J Immunother. 2014; 37:257–66. https://doi. org/10.1097/CJI.0000000000000033.

45. Castells A, Ino Y, Louis DN, Ramesh V, Gusella JF, Rustgi AK. Mapping of a target region of allelic loss to a 0.5-cM interval on chromosome 22q13 in human colorectal cancer. Gastroenterology. 1999; 117:831–7.

46. Cliby W, Ritland S, Hartmann L, Dodson M, Halling KC, Keeney G, Podratz KC, Jenkins RB. Human epithelial ovarian cancer allelotype. Cancer Res. 1993; 53:2393–8.

47. Zweemer RP, Ryan A, Snijders AM, Hermsen MA, Meijer GA, Beller U, Menko FH, Jacobs IJ, Baak JP, Verheijen RH, Kenemans P, van Diest PJ. Comparative genomic hybridization of microdissected familial ovarian carcinoma: two deleted regions on chromosome 15q not previously identified in sporadic ovarian carcinoma. Lab Invest. 2001; 81:1363–70.

48. Chong GO, Jeon HS, Han HS, Son JW, Lee YH, Hong DG, Lee YS, Cho YL. Differential MicroRNA Expression Profiles in Primary and Recurrent Epithelial Ovarian Cancer. Anticancer Res. 2015; 35:2611–7.

49. Palmero EI, Galvao HC, Fernandes GC, Paula AE, Oliveira JC, Souza CP, Andrade CE, Romagnolo LG, Volc S, C Neto M, Sabato C, Grasel R, Mauad E, et al. Oncogenetics service and the Brazilian public health system: the experience of a reference Cancer Hospital. Genet Mol Biol. 2016; 39:168–77. https://doi.org/10.1590/1678-4685-GMB-2014-0364.

50. van Beers EH, Joosse SA, Ligtenberg MJ, Fles R, Hogervorst FB, Verhoef S, Nederlof PM. A multiplex PCR predictor for aCGH success of FFPE samples. Br J Cancer. 2006; 94:333–7. https://doi.org/10.1038/sj.bjc.6602889.

6. ESTUDO DE VARIANTES GERMINATIVAS PATOGÊNICAS NAS AMOSTRAS *BRCA1*/*BRCA2*-NEGATIVAS, ATRAVÉS DE SEQUENCIAMENTO EXÔMICO

Os resultados apresentados nessa seção foram divididos em duas etapas, sendo elas:

6.1. - Resultados submetidos para publicação no periódico "*Cancer"*, com fator de impacto: 6,53. O arquivo PDF do artigo publicado na revista científica encontra-se anexado à esta tese (Anexo III).

6.2. - Resultados serão submetidos para publicação no periódico "*Human Mutation*", com fator de impacto: 4,45.

6.1. Whole-exome sequencing of Brazilian non-*BRCA1/BRCA2* mutation carrier cases at high-risk for hereditary breast/ovarian cancer

Felicio PS, Grasel RS, Campacci N, Galvão HCR; Torrezan GT; de Paula AE; Sabato CS; Fernandes GC; Souza CP, Michelli RD, Andrade CE, Revil T, Ragoussis J, Couch FJ, Hart SN, Reis RM, Tonin PN, Melendez ME, Carraro DM, Palmero EI.

Abstract

The current study aimed to identify, through whole-exome sequencing (WES) of germline/constitutional DNA, new breast and/or ovarian cancer predisposition genes in Brazilian families at high-risk for hereditary breast/ovarian cancer (HBOC). We performed WES in 52 non-*BRCA1/BRCA2* mutation carrier women with high-risk for HBOC. All identified variants were classified using information from population and disease specific databases, as well as *in silico* prediction tools and the ACMG criteria. The majority of patients (n=32, 61.5%) with breast/ovarian cancers were diagnosed at early ages (≤45 years old). Moreover, 87% (n=45) and 31% (n=16) of women reported the presence of a family history of breast or ovarian cancers, respectively. The WES analysis showed the presence of a total of 53 unique loss of function variants and 128 unique rare missense variants. Of these, 23 genes were described as cancer hallmarks by the COSMIC database, including *ATM*, *CHEK2*, *PMS2*, *KRAS*, *KIT* and *EGFR*. Furthermore, variants in the RAD gene family were observed, such as *RAD50*, *RAD51C* and *RAD54L*. In conclusion this is the largest Brazilian WES study involving families at high-risk

for HBOC that has brought new insights of the role of potentially new genetic risk factors in Brazilian HBOC.

Keywords: BRCAX, non-BRCA, whole-exome sequencing, hereditary breast and ovarian cancer predisposition syndrome.

INTRODUCTION

Germline variants in *BRCA1/BRCA2* are responsible for approximately 25% of the familial breast cancer (BC) and ovarian cancer (OC) cases, and predispose to the hereditary breast and ovarian cancer (HBOC) syndrome worldwide[1,2]. In Brazil, our group recently reported similar findings, with 21.5% of the 349 index cases with clinical criteria for HBOC syndrome harboring *BRCA1/BRCA2* germline variants[3]. Genomic advances, such as next generation DNA sequencing platforms, allows the analysis of gene panels and the subsequent association of other high and moderated risk genes for HBOC with hereditary BC and OC development. These genes include, among others, *ATM*, *BRIP1*, *CDH1*, *PALB2*, *PTEN*, *RAD51C*, *STK11* and *TP53*[4]. However, for a large proportion of HBOC families (50-80%)[2,5] the genetic cause associated with the BC and OC family history is unknown. A deeper genomic analysis, such as using whole-exome sequencing (WES) followed by bioinformatic analysis of rare variants might reveal new cancer predisposing alleles.

WES can be effective in the diagnosis of individuals for whom the traditional approaches were not conclusive. Besides, it can decreases the exhausting diagnostic odyssey in families with rare or uncharacterized diseases[6]. Additionally, WES can provide opportunities for personalized health-care strategies, such as prevention or early detection of diseases[7]. Several researchers support the idea that WES is the most appropriate tool for identifying genetic familial syndromes. According to Snape *et al.*[8], the analysis of WES offers the potential to perform a strategy not based only on a candidate gene, but on the information available throughout the WES, in search for possible variants with potential significance for the disease ("agnostic approach"). Moreover, as reported by Cooper *et al.*[9], 85% of the mutations relevant to human are located in the coding region or in canonical splice sites. Therefore, WES could be a promising strategy for identifying new genes associated with an increased risk of BC and OC in high-risk families that do not harbor known pathogenic germline variants in *BRCA1* and *BRCA2*[10].

The aim of the current study was to perform WES in 52 unrelated Brazilian women with high-risk for BC and OC, previously tested negative for pathogenic *BRCA1/BRCA2* germline variants in order to identify driver genes of HBOC.

MATERIALS AND METHODS

Ethical criteria

The present study was approved by Barretos Cancer Hospital (BCH) ethics committee (approval number: 40814115.4.0000.5437), and all participants gave their consent to participate.

Patients

Fifty-two unrelated Brazilian women at-risk for HBOC attended at the Oncogenetics Department of BCH[11] were included. All cases had a personal and family history of BC and/or OC. Analysis of the presence of germline variants in *BRCA1/BRCA2/TP53* genes was conducted at the Center of Molecular Diagnosis of BCH as part of routine care through Sanger/NGS sequencing followed by rearrangement analysis through MLPA (Multiplex Ligation-dependent Probe Amplification Analysis), as described elsewhere by Fernandes *et al*. (2016)[3]. Clinical information was obtained through detailed review of the patient´s clinical chart. For family history of cancer data, all pedigrees were studied.

DNA isolation and quantification for WES

Genomic DNA was isolated from peripheral blood lymphocytes using the QIAamp DNA Blood Mini Kit (Qiagen) following the manufacturer's instructions. DNA concentration was determined using Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific).

Whole-exome sequencing

For the exome library preparation, 50ng of constitutional DNA of each sample was used. The kit used was the Nextera Rapid Capture Expanded Exome (Illumina), according to the manufacturer's recommendations. Quantification of the enriched library was performed with Qubit fluorometer (Thermo Fisher Scientific) and library size distribution was measured with Agilent Bioanalyzer 2100 (Agilent Genomics). Quantified DNA library was loaded on flow

cell for subsequent cluster generation. Samples were paired-end sequenced on Illumina NextSeq 500 High Output Kit - 300 cycles (Illumina).

Whole-exome sequencing analysis

Briefly, reads were quality trimmed using the Trimmomatic v0.33[12], and then aligned with the genome of reference (UCSC GRCh37/hg19) using the Burrows-Wheeler Aligner (BWA) v0.7.5a. PCR duplicates were removed using Picard v1.106 and BAM files were processed using the Genome Analysis Toolkit (GATK) v2.7.2 software. Realignment and search for indels were performed using GATK HaplotypeCaller and annotated using snpEFF v4.3 and SnpSift[13]. A GEMINI v.0.19.1 database was created[14], and variants selected per functional rules. Additionally, variants described by snpEFF/GEMINI as "low-impact" were removed since they are assumed to have benign effects on DNA or protein behavior.

The analysis workflow is illustrated in Figure 1. For a function-based prioritization, variants leading to loss of function ("high-impact" variants: frameshift, nonsense, and canonical splice site variants) and missense variants (classified as "medium-impact" variants) were selected. For quality filtering, variants with vertical coverage ≥10x and variant allele frequency (VAF) ≥0.25 were selected. Next, a total of 2,319 cancer-associated genes were analyzed (described below). Variants present in the population database Genome Aggregation Database (gnomAD[15]), with a frequency ≤2% (minor allele frequency [MAF] ≤0.02) were maintained. Furthermore, a recently publicly available Brazilian database of WES from 609 healthy individuals (AbraOM—Brazilian genomic variants[16]) was also used for manually excluding population-specific variants (MAF ≤0.02 were maintained). Loss of function variants were manually examined with Integrative Genomics Viewer (IGV)[17] to remove possible artifacts.

Cancer gene reference lists

Three databases were used to generate a candidate list of genes previously reported associated with any type of cancer, namely: i) The Cancer Gene Census v.86, a set of 719 genes manually curated by the Sanger Institute[18], ii) a query of DISEASES[19], a database of disease-gene associations based largely on text-mining approaches, and, iii) UniprotKB[20], a manually curated database of protein functions (using the keyword-terms "cancer", "tumor-suppressor gene", "proto-oncogene" and "oncogene"). From these databases, a reference list

of 2,319 genes was generated for prioritizing and characterizing gene variants. Detailed information about these genes is available in Table S1.

RESULTS

1. Clinico-pathological characterization

Of the 52 Brazilian women analyzed, 41 had BC, nine had OC and two were initially diagnosed with OC which was followed by a BC diagnosis. Detailed clinico-pathological features are specified in Table S2. Briefly, the average age of cancer diagnosis was 41.1 years, ranging from 20 to 60 years. The majority of BC cases had invasive ductal carcinoma (85.4%), and were estrogen and progesterone positive (70.3% and 60.5%, respectively) and HER2 negative (77.8%) tumors. Most of the BC cases were molecularly classified as luminal type tumors (73.0%), seven were triple negative tumors (18.9%) and three (8.1%) were tumors overexpressing HER2. The majority of women with OC developed high-grade serous adenocarcinoma subtype tumors (90.9%).

2. Family history

The family history of the patients included in this study is depicted in Table S3. Besides BC and OC, other cancers associated with HBOC spectrum, such as cancers of the prostate (n=14 relatives, 10 families) and pancreas (n=4 relatives, 4 families) were observed. BC cases were observed in first degree relatives in 56% of the families (n=29). In addition, the presence of OC among first degree relatives was observed in four families, and in three of these four families the proband was also diagnosed with OC.

3. Germline variants by whole-exome sequencing

The WES identified a total of 2,536,915 variants in the 52 cases (Figure 1). To narrow down the analysis, a reference list of 2,319 candidate genes previously associated with cancer was used. After applying the depicted workflow (Figure 1), a total of 3,027 variants were selected for further bioinformatic analyses. Next, variants found in ≥25% of our cohort (13 patients) were excluded (totaling 183 variants) as they likely represent sequencing and/or annotation artefacts. Therefore, a total of 2,844 variants were analyzed in this study. Of these, 75 (2.7%) were classified as loss of function (LoF) and 2,769 (97.3%) were missense/inframe variants.

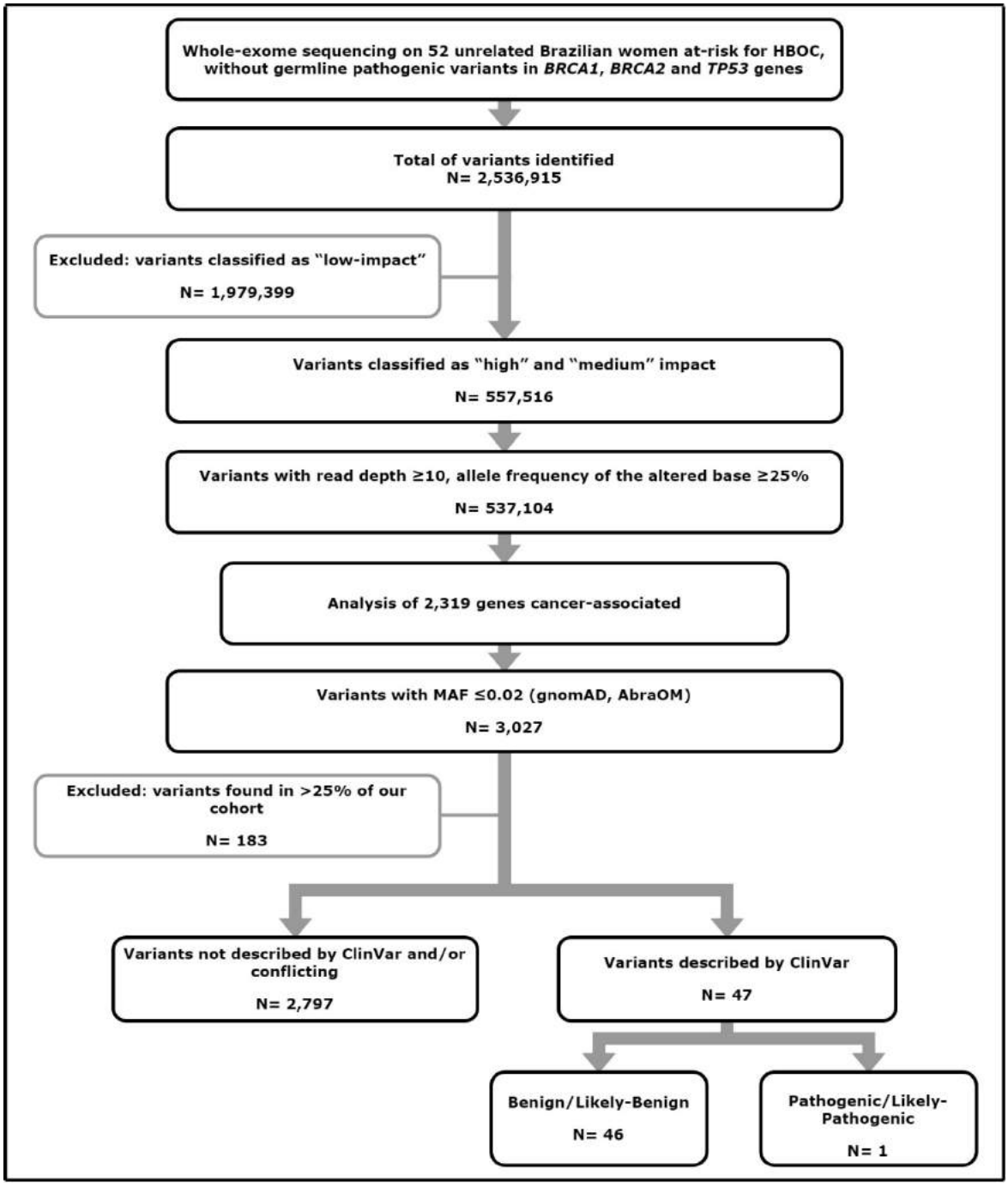


Figure 1 - Variants selection workflow. Whole-exome sequencing data from 52 unrelated Brazilian women at-risk for HBOC, without germline pathogenic variants in *BRCA1*, *BRCA2* and *TP53* genes. Variants classified as "high-impact" and "medium impact" by snpEFF/GEMINI were prioritized. Then, variants with base coverage ≥10x and variant allele frequency (VAF) ≥0.25 were selected, and those present in population databases with frequency ≤2% (MAF ≤0.02) were analyzed. The variants were also separated accordingly to ClinVar classification.

The frequency of variants in each case is detailed in Figure 2. Overall, an average of 55 variants was observed, ranging from 15 to 131 variants per patient. In addition, when the variants were grouped by the impact, an average of 1 LoF variant and 53 missense/inframe variants were identified in each patient. Moreover, germline variants described by ClinVar as benign/likely-benign (B/LB) and pathogenic/likely-pathogenic (P/LP) were analyzed separately. Of these, 46 variants were described as B/LB by ClinVar and, only one variant with LoF classified as LP/P (without conflicting interpretations) by ClinVar. The frameshift variant classified as pathogenic by ClinVar (*PMS2*: c.2182_2184delinsG) was identified in patient 306, who had melanoma at 26 years of age and BC at 36 years of age.

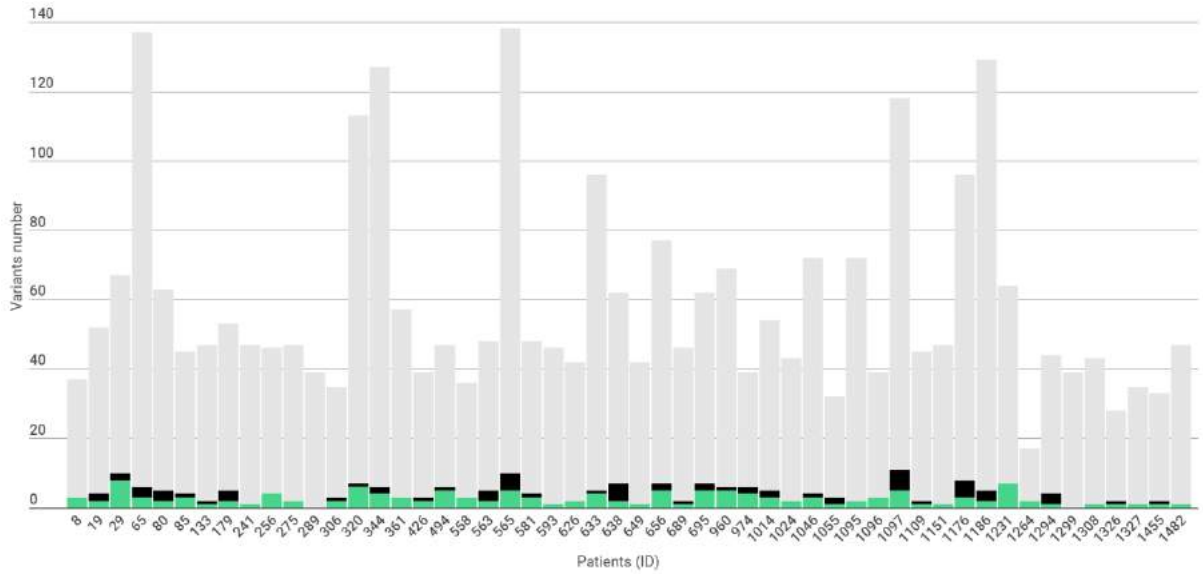


Figure 2 – Total number of variants per patient. In gray: total number of variants in cancer-associated genes, before applying the ClinVar classification (pathogenic/likely-pathogenic, variants with conflicting interpretations and benign/likely-benign) and in silico tools (CADD, REVEL and M-CAP); in black: total number of variants classified as LoF with conflicting interpretations, VUS or not classified by ClinVar; in green: total number of rare missense variants with conflicting interpretations, VUS or not classified by ClinVar and predicted as pathogenic by in silico analysis (CADD ≥20, REVEL ≥0.5 and M-CAP ≥0.025).

3.1. Germline variants with conflicting interpretations, VUS or not classified by ClinVar

Variants with discordance and/or not reported in the ClinVar database were evaluated separately. A total of 2,797 variants were identified in this part of the analysis, including 68 LoF and 2,729 missense/inframe variants. An average of 52 variants in each case was

observed, ranging from 13 to 131 variants. The frequency and the type of variants found in each case, after applying the filters are detailed in Figure 2.

3.1.1. LoF variants

After applying all filters described above, a total of 52 unique LoF variants (with conflicting interpretations, VUS or not classified by ClinVar) were identified (19 frameshift, 23 nonsense and 10 splice acceptor/donor variant). All LoF variants identified in genes associated with cancer are detailed in Table 1 and Figure S1.

Table 1 - Rare variants identified in cancer-associated genes.

Sample ID	Personal history of cancer	BC - Family history	OC – Family history	Gene	Variant type	RS	HGVS - c.	HGVS - p.	Transcript	ClinVar	Hallmark (COSMIC)
19	BC, 27	Yes	No	*RAD54B*	frameshift	rs142650111	c.2733del	p.T911Yfs	ENST00000336148	Not reported	No
				STAP2	nonsense	rs79657645	c.507C>G	p.Y169*	ENST00000594605	Not reported	No
29	OC, 42 / BC, 53	Yes	Yes	*KMT2E*	splice donor	rs74959149	c.1759+1G>A	.	ENST00000334914	Not reported	No
				TPO	nonsense	rs150762359	c.2450G>A	p.W817*	ENST00000422464	Not reported	No
65	BC, 36	Yes	No	*EPHB2*	nonsense	rs76826147	c.3055A>T	p.K1019*	ENST00000400191	Not-provided	No
				FLCN	nonsense	rs201282009	c.1029A>C	p.T343*	ENST00000389169	Not reported	No
				LCP1	nonsense	Not reported	c.1122C>A	p.Y374*	ENST00000323076	Not reported	No
80	BC, 43	Yes	No	*BCR*	frameshift	rs372013175	c.3275_3278dup	p.V1094Rfs*17	ENST00000305877	Not reported	No
				CCND3	nonsense	rs33966734	c.379G>T	p.E127*	ENST00000510503	Not reported	No
				FANCA	splice donor	rs55873203	c.435+2dup	.	ENST00000561722	Not reported	No
85	BC, 51	Yes	No	*CCND3*	nonsense	rs33966734	c.379G>T	p.E127*	ENST00000510503	Not reported	No
133	BC, 46	Yes	Yes	*DROSHA*	nonsense	Not reported	c.1498G>T	p.E500*	ENST00000344624	Not reported	Yes
179	BC, 47	Yes	No	*BCR*	frameshift	rs372013175	c.3275_3278dup	p.V1094Rfs*17	ENST00000305877	Not reported	No
				CCND3	nonsense	rs33966734	c.379G>T	p.E127*	ENST00000510503	Not reported	No
				STAP2	nonsense	rs79657645	c.507C>G	p.Y169*	ENST00000594605	Not reported	No
306	Melanoma, 26 / BC, 36	Yes	No	*PMS2*	frameshift	rs1554294508	c.2182_2184delinsG	p.T728Afs	ENST00000265849	Pathogenic	Yes
320	OC, 53	No	Yes	*NHEJ1*	splice acceptor	rs10498064	c.191-2A>T	.	ENST00000483627	Not reported	No
344	OC, 47	No	Yes	*PLK2*	frameshift	Not reported	c.1004dupT	p.L335Ffs*12	ENST00000274289	Not reported	No
				USP6	frameshift	rs560633910	c.2079_2080del	p.C694Wfs*8	ENST00000250066	Not reported	No
494	BC, 33	Yes	Yes	*IGF2*	nonsense	rs200441006	c.97C>T	p.Q33*	ENST00000434045	Not reported	No

563	BC, 39	Yes	No	*CTNNA1*	frameshift	Not reported	c.1206_1207insCC	p.V403Pfs*3	ENST00000302763	Not reported	No
				MITF	nonsense	Not reported	c.1A>G	p.M1V	ENST00000448226	Not reported	No
				TCF3	splice acceptor	rs41275842	c.52-2A>G	.	ENST00000453954	Not reported	No
565	OC, 43	No	Yes	*BCAR1*	splice donor	rs74024754	c.12+2T>C	.	ENST00000535626	Not reported	No
				KAT5	splice donor	rs112903919	c.70+1G>C	.	ENST00000533596	Not reported	No
				NEIL1	splice donor	rs5745908	c.434+2T>C	.	ENST00000355059	Not reported	No
				SLC34A2	splice acceptor	rs199782502	c.113-2A>G	.	ENST00000513204	Not reported	Yes
				ZNF429	frameshift	rs199679715	c.1580del	p.H527Lfs*157	ENST00000358491	Not reported	No
581	BC, 46	Yes	No	*NBPF3*	nonsense	Not reported	c.2T>C	p.M1V	ENST00000318249	Not reported	No
633	BBC, 38	Yes	No	*HERPUD1*	frameshift	rs138682179	c.582_583del	p.F195Cfs*32	ENST00000563911	Not reported	No
638	BC, 42	Yes	No	*AIM2*	nonsense	rs74689714	c.1029+1T>G	.	ENST00000368130	Not reported	No
				CNBD1	nonsense	rs78702891	c.2T>C	p.M1T	ENST00000518476	Not reported	No
				POLQ	frameshift	rs546221341	c.4262_4268del	p.I1421Rfs*8	ENST00000264233	Not reported	No
				PTPRH	nonsense	rs147881000	c.2125C>T	p.Q709*	ENST00000263434	Not reported	No
				RGPD3	nonsense	rs569801331	c.93T>A	p.Y31*	ENST00000304514	Not reported	No
656	OC, 41 / BC: 48	No	Yes	*KAT5*	splice donor	rs112903919	c.70+1G>C	.	ENST00000533596	Not reported	No
				RAD51C	frameshift	Not reported	c.890_899delTTGTTCCTGC	p.(Leu297HisfsTer2)	ENST00000337432	Not reported	No
689	BBC, 47	Yes	No	*POLQ*	frameshift	rs546221341	c.4262_4268del	p.I1421Rfs*8	ENST00000264233	Not reported	No
695	OC, 21	Yes	No	*MED28*	nonsense	Not reported	c.52C>T	p.Q18*	ENST00000237380	Not reported	No
				PTTG2	frameshift	rs200376306	c.554_555del	p.H185Lfs*16	ENST00000504686	Not reported	No
960	BBC, 59 and 70	Yes	No	*KMT2E*	splice donor	rs74959149	c.1759+1G>A	.	ENST00000334914	Not reported	No
974	BC, 46	Yes	No	*BCR*	frameshift	rs372013175	c.3275_3278dup	p.V1094Rfs*17	ENST00000305877	Not reported	No
				FAN1	frameshift	Not reported	c.357_358delGG	p.(Glu120SerfsTer10)	ENST00000362065	Not reported	No

1014	BC, 42	Yes	No	*CCND3*	nonsense	rs33966734	c.379G>T	p.E127*	ENST00000510503	Not reported	No
				KMT2C	nonsense	rs58528565	c.2961C>G	p.Y987*	ENST00000262189	Not reported	No
1046	BC, 37	Yes	No	*PDE4DIP*	Nonsense	rs587666271	c.287G>T	p.C96F	ENST00000491426	Not reported	No
1055	OC, 57	Yes	No	*PTCH1*	Nonsense	rs755103500	c.1A>G	p.M1V	ENST00000375274	Not reported	No
				RIPK1	Frameshift	Not reported	c.1802_1805delGTGC	p.C601Sfs*23	ENST00000259808	Not reported	No
1097	OC, 46	Yes	No	*AFF3*	Frameshift	rs201754690	c.1788del	p.D598Tfs*48	ENST00000317233	Not reported	Yes
				CTNNA2	splice acceptor	rs75100314	c.103-1G>A	.	ENST00000409266	Not reported	No
				ENO2	Nonsense	Not reported	c.1235+2_1235+5dupTGAG	p.I413*fs*1	ENST00000229277	Not reported	No
				MELK	Nonsense	rs770129579	c.693G>A	p.W231*	ENST00000298048	Not reported	No
				PUM1	Nonsense	rs116992433	c.229T>C	.	ENST00000530669	Not reported	No
				TDG	splice donor	rs765686214	c.1090_1090+1insTTGAGAGC	p.V367Lfs*6	ENST00000392872	Not reported	No
1109	BC, 21	No	Yes	*TCF3*	splice acceptor	rs41275842	c.52-2A>G	.	ENST00000453954	Not reported	No
1176	BC, 44	No	Yes	*CDK20*	Frameshift	Not reported	c.516dupC	p.E173Rfs*10	ENST00000325303	Not reported	No
				IGF2	Nonsense	rs200441006	c.97C>T	p.Q33*	ENST00000434045	Not reported	No
				PTTG2	Frameshift	rs200376306	c.554_555del	p.H185Lfs*16	ENST00000504686	Not reported	No
				RFC3	Nonsense	Not reported	c.*1C>A	.	ENST00000380071	Not reported	No
				TCF3	splice acceptor	rs41275842	c.52-2A>G	.	ENST00000453954	Not reported	No
1186	OC, 20	No	No	*POLM*	Frameshift	rs28382660	c.1191del	p.K398Sfs*25	ENST00000395831	Not reported	No
				RBM14	Frameshift	rs747273987	c.373del	p.S125Lfs*97	ENST00000409372	Not reported	No
				TSC2	Frameshift	Not reported	c.5068+2_5068+6dupTAGGG	.	ENST00000219476	Not reported	No
1294	BC, 26	Yes	No	*HERPUD1*	Frameshift	rs138682179	c.582_583del	.	ENST00000563911	Not reported	No
				KMT2C	Nonsense	rs58528565	c.2961C>G	p.Y987*	ENST00000262189	Not reported	No
				RGPD3	Frameshift	rs771658117	c.3826_3830del	p.S1276Cfs*9	ENST00000304514	Not reported	No
1326	BC, 35	Yes	No	*RGPD3*	Frameshift	rs771658117	c.3826_3830del	p.S1276Cfs*9	ENST00000304514	Not reported	No

1455	BC, 31	Yes	No	*SP100*	Frameshift	rs762056038	c.2282dup	p.K762Efs*21	ENST00000264052	Not reported	No

Subtitles: BBC: bilateral breast cancer; BC: breast cancer; OC: ovarian cancer.

The most frequently mutated gene with LoF variants was *CCND3*. A unique nonsense variant identified in *CCND3* gene (c.379G>T) was observed in four unrelated women (patients ID: 80, 85, 179 and 1014) with BC (average age at diagnosis: 46 years). All of *CCND3* variant carriers reported a family history of BC. The presence of this mutation in the four families from each variant carrier was confirmed by Sanger sequencing.

Variants were identified in in genes described by the COSMIC database as "hallmarks of cancer", such as *AFF3*, *DROSHA* and *SLC34A2* (Figure 3). The frameshift variant identified in the *AFF3* (c.1788del) has not been described by ClinVar. The patient (ID 1097) identified with this variant had OC at 46 years of age and reported a family history of BC. The nonsense variant in *DROSHA* (c.1498G>T), also not reported by ClinVar, was identified in a patient (ID 133) diagnosed with BC at 46 years of age who also reported a family history of BC. The splice acceptor variant in *SLC34A2* (c.113-2A>G), not described by ClinVar, was identified in a patient (ID 565) who reported a personal/family history of OC.

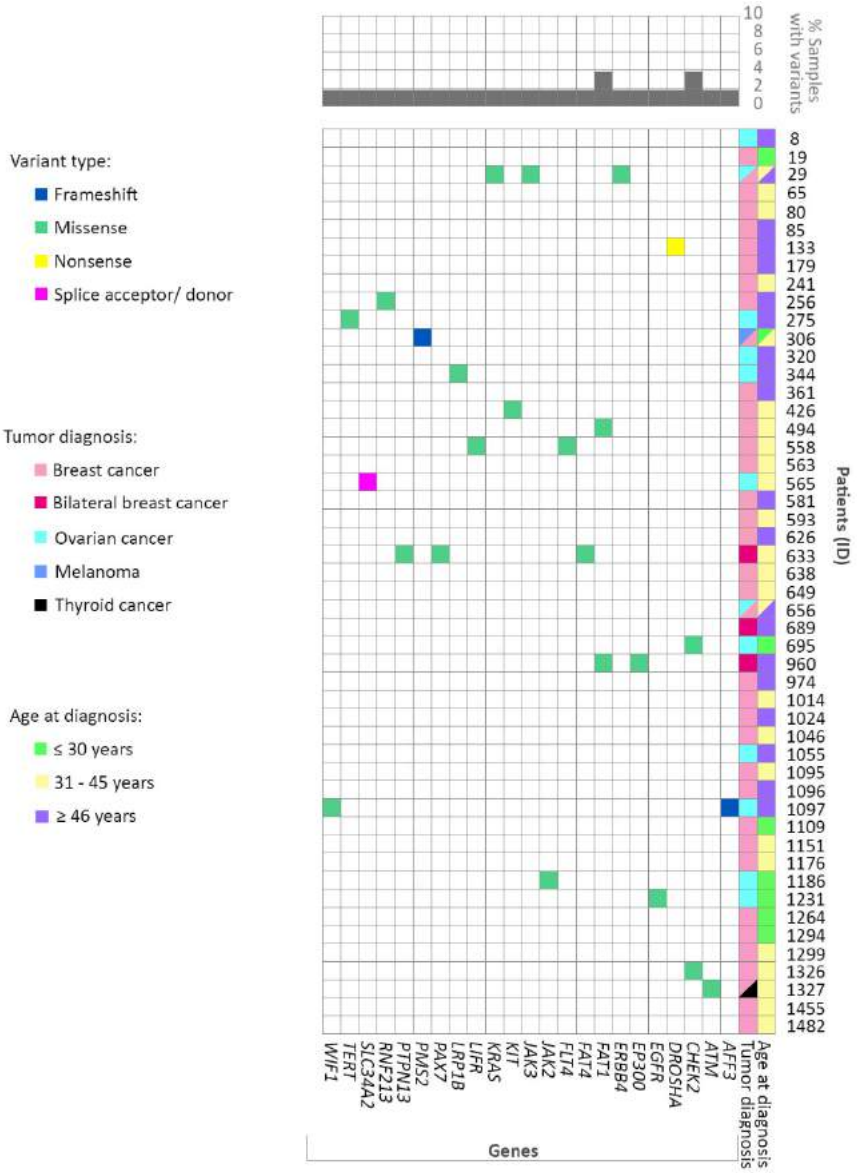


Figure 3 – Rare variants identified involving genes classified as hallmarks of cancer by COSMIC database. In blue: frameshift variants with conflicting interpretations, VUS or not classified by ClinVar; in yellow: nonsense variants with conflicting interpretations, VUS or not classified by ClinVar; in purple: splice acceptor/donor variants with conflicting interpretations, VUS or not classified by ClinVar; in green: missense variants with conflicting interpretations, VUS or not classified by ClinVar and predicted as pathogenic by in silico analysis (CADD ≥20, REVEL ≥0.5 and M-CAP ≥0.025). Information about tumor diagnosis (in light pink: breast cancer; in dark pink: bilateral breast cancer; in greenish-blue: ovarian cancer; in light blue: melanoma; in black: thyroid cancer) and age at diagnosis (in fluorescent green: diagnosis ≤ 30 years of age; in light yellow: 31 to 45 years of age; in purple: ≥ 46 years of age) are represented.

Besides, identifying variants in genes known associated with other types of hereditary/familial tumors, LoF variants in other cancer associated genes such as *FANCA*, *FLCN*, *POLQ*, *PTCH1*, *RAD51C*, *RAD54B* and *TSC2* were also identified in our high-risk cases. The variant identified in *POLQ*, (c.4262_4268del), was found in two unrelated BC patients. One of them (patient ID 689) had bilateral BC at 47 years of age and reported a family history of leukemia, BC and colorectal tumors. The other patient (ID 638) with the same *POLQ* variant, had BC at 42 years of age and described a family history BC and gastric tumors. The frameshift variant identified in *RAD51C* (c.890_899delTTGTTCCTGC) was observed in patient 656, who had OC at 41 years of age and BC at 48 years of age, and reported a second degree relative with OC. A frameshift variant in another gene in the *RAD*-family, the *RAD54B* (c.2733del), was identified in a patient (ID 19) with BC at 27 years of age, who reported two relatives with BC (before 45 years of age) in their family history. The frameshift variant identified in the *TSC2* (c.5068+2_5068+6dupTAGGG) was observed in a patient (1186) who had OC at 20 years of age and, one year later, was diagnosed with an appendiceal neuroendocrine tumor. This patient reported a second-degree relative uterine cancer. The nonsense variants identified in *FLCN* (c.1029A>C) was observed in patient 65, who had a personal/familial of BC. The variant identified in *PTCH1* (c.1A>C) was detected in patient 1055, who was diagnosed with OC at 57 years of age and reported the presence of BC in their family history. The splice donor variant involved in the *FANCA* gene (c.435+2dup) was observed in patient 80. This woman had BC at 43 years of age and reported six relatives with BC in her family history. The significance of any of above described LoF variants are unknown as none of them have been reported in the ClinVar database.

3.1.2. Missense variants

To further evaluate the missense variants, three bioinformatic tools were selected and applied in the analysis of our variants, namely CADD (score ≥20), REVEL (score ≥0.5) and M-CAP (score ≥0.025). The number of variants de-prioritized by each tool is detailed in Figure S2. As a result of this filtering step, a total of 136 rare missense variants were selected (128 unique variants) for further evaluation. There was an average of three variants per patient, ranging from 1 to 8 variants. All selected missense variants that were considered pathogenic by the three bioinformatic tools, together with information regarding the patient´s age and cancer type at diagnosis are detailed in Table 2 and Figure S1.

Table 2 - Rare missense variants identified in cancer-associated genes

Sample ID	Personal history of cancer	BC - Family history	OC - Family history	Gene	RS	HGVS - c.	HGVS - p.	Transcript	ClinVar	CADD scaled	REVEL score	M-CAP score	Hallmark (COSMIC)
8	BBC, 51 and 52	Yes	No	*CUL7*	rs147493246	c.4762C>A	p.L1588I	ENST00000265348	Uncertain, Benign	32.0	0.592	Not reported	No
				CUZD1	rs36212072	c.854T>C	p.I285T	ENST00000368904	Not reported	22.9	0.572	Not reported	No
				MC1R	rs1805009	c.880G>C	p.D294H	ENST00000555147	Likely-benign, Benign, Pathogenic	28.3	0.608	Not reported	No
19	BC, 27	Yes	No	*COL2A1*	rs764878166	c.371G>A	p.G124D	ENST00000337299	Not reported	21.3	0.653	0.1434	No
				MME	rs138218277	c.1040A>G	p.Y347C	ENST00000360490	Not reported	23.6	0.556	0.0409	No
29	OC, 42 / BC,53	Yes	Yes	*A1CF*	None	c.107G>A	p.G36E	ENST00000282641	Not reported	32.0	0.611	0.1117	No
				CHD1L	None	c.413T>C	p.L138P	ENST00000361293	Not reported	28.2	0.899	0.2244	No
				ERBB4	rs770938636	c.3446G>T	p.G1149V	ENST00000342788	Not reported	33.0	0.627	0.1794	Yes
				JAK3	rs149452625	c.2773C>A	p.R925S	ENST00000458235	Not reported	23.4	0.52	0.3413	Yes
				KAT6B	rs746792857	c.2495A>T	p.N832I	ENST00000287239	Not reported	25.4	0.6	0.1739	No
				KRAS	None	c.461A>G	p.D154G	ENST00000256078	Not reported	28.6	0.751	0.1119	Yes
				MYO5A	None	c.4435C>G	p.P1479A	ENST00000356338	Not reported	26.7	0.557	0.100	No
				PMS1	rs1145232	c.856G>A	p.G286R	ENST00000409593	Not-provided, Likely-benign	25.9	0.779	Not reported	No
65	BC, 36	Yes	No	*IFRD2*	rs2229648	c.1811G>A	p.R604Q	ENST00000336089	Not reported	35.0	0.634	Not reported	No
				NDRG2	rs36007455	c.91A>T	p.T31S	ENST00000298684	Not reported	26.8	0.503	Not reported	No
				PMS1	rs147566508	c.1625G>A	p.R542H	ENST00000409593	Not reported	27.0	0.69	0.1809	No
80	BC, 43	Yes	No	*KDM5A*	None	c.1049A>G	p.E350G	ENST00000382815	Not reported	31.0	0.725	0.1789	No
				LEPROTL1	None	c.193G>A	p.A65T	ENST00000321250	Not reported	28.9	0.546	0.0329	No
85	BC, 51	Yes	No	*MLH1*	rs63751448	c.794G>A	p.R265H	ENST00000231790	Uncertain	34.0	0.87	0.6093	No
				UNG	rs151095402	c.262C>T	p.R88C	ENST00000242576	Not reported	35.0	0.545	0.1911	No
133	BC, 46	Yes	Yes	*KANK1*	rs61737971	c.149A>T	p.D50V	ENST00000382297	Not reported	26.0	0.54	0.0529	No

179	BC, 47	Yes	No	*ERCC8*	rs61754098	c.839C>A	p.T280K	ENST00000265038	Uncertain, Likely-benign	24.4	0.565	0.0680	No
				PTPRK	rs771714409	c.4292G>A	p.R1431Q	ENST00000368207	Not reported	29.6	0.517	0.0649	No
241	BC, 45	Yes	No	*CNTNAP2*	rs139694086	c.400T>G	p.W134G	ENST00000361727	Uncertain	23.6	0.897	0.1300	No
256	BC, 47	Yes	Yes	*RFC1*	rs28903096	c.2014G>A	p.V672M	ENST00000349703	Not reported	34.0	0.776	0.3881	No
				RNF213	rs61741961	c.8249G>T	p.W2750L	ENST00000336301	Not reported	28.0	0.602	Not reported	Yes
				RPN1	rs138936459	c.712C>T	p.H238Y	ENST00000296255	Not reported	29.5	0.765	Not reported	No
				TES	rs754828812	c.1166T>C	p.F389S	ENST00000358204	Not reported	32.0	0.969	0.4575	No
275	OC, 60	Yes	No	*CCNG1*	rs747511557	c.441G>C	p.W147C	ENST00000340828	Not reported	22.5	0.581	0.0262	No
				TERT	rs370686937	c.3332C>T	p.T1111M	ENST00000310581	Not reported	24.3	0.56	0.8051	Yes
306	Melanoma, 26 / BC, 36	Yes	No	*MC1R*	rs1805009	c.880G>C	p.D294H	ENST00000555147	Likely-benign, Benign, Pathogenic	28.3	0.608	Not reported	No
				WFDC2	rs777713492	c.97C>T	p.P33S	ENST00000339946	Not reported	25.2	0.639	0.1851	No
320	OC, 53	No	Yes	*DMBT1*	rs377585441	c.2018G>A	p.G673E	ENST00000338354	Not reported	20.5	0.604	0.0556	No
				FKBP9	rs150348129	c.940G>A	p.V314M	ENST00000242209	Not reported	31.0	0.736	Not reported	No
				GATA2	rs370164300	c.1348G>A	p.G450R	ENST00000341105	Not reported	27.2	0.644	0.5836	No
				KRT19	rs117671585	c.560G>A	p.R187H	ENST00000361566	Not reported	33.0	0.678	0.0933	No
				NEK3	rs201880627	c.29T>A	p.I10N	ENST00000339406	Not reported	32.0	0.636	0.2419	No
				RAD54L	rs28363218	c.604C>T	p.R202C	ENST00000371975	Not reported	34.0	0.714	0.1532	No
344	OC, 47	No	Yes	*CSMD3*	rs145027071	c.9271T>C	p.S3091P	ENST00000297405	Not reported	25.2	0.504	Not reported	No
				LRP1B	rs72899872	c.9532G>A	p.A3178T	ENST00000389484	Not reported	33.0	0.734	Not reported	Yes
				PAK4	rs146583353	c.61G>A	p.V21M	ENST00000321944	Not reported	29.0	0.766	0.3938	No
				VRK1	rs146113610	c.683C>T	p.T228M	ENST00000216639	Not reported	34.0	0.513	0.1720	No
361	BC, 56	Yes	No	*BIRC6*	None	c.9838C>T	p.L3280F	ENST00000421745	Not reported	29.7	0.523	0.195194712941	No
				DNMT3A	rs143730975	c.89A>C	p.E30A	ENST00000264709	Not-provided	23.1	0.501	Not reported	No
				MSH2	rs750746034	c.80C>T	p.P27L	ENST00000233146	Uncertain	34.0	0.683	0.586681149373	No
426	BC, 38	Yes	No	*FAN1*	rs148404807	c.149T>G	p.M50R	ENST00000362065	Not reported	28.4	0.761	0.122156593397	No
				KIT	None	c.50T>C	p.L17P	ENST00000288135	Not reported	21.3	0.535	0.391470428989	Yes

494	BC, 33	Yes	Yes	*ABCG2*	None	c.802G>A	p.G268R	ENST00000237612	Not reported	33.0	0.744	0.0335842111063	No
				FAT1	rs200149452	c.7957G>A	p.G2653S	ENST00000441802	Not reported	24.1	0.69	0.110363283318	Yes
				TFRC	rs866295059	c.2123C>T	p.T708M	ENST00000360110	Not reported	25.0	0.563	0.12209150802	No
				UFL1	rs199880163	c.2114C>G	p.P705R	ENST00000369278	Not reported	28.3	0.622	0.0783739696574	No
				WWOX	rs781063964	c.253T>G	p.Y85D	ENST00000355860	Not reported	23.9	0.894	0.276836072671	No
558	BC, 37	Yes	Yes	*FLT4*	rs149033942	c.3023C>T	p.P1008L	ENST00000261937	Not reported	23.5	0.651	0.1684	Yes
				LIFR	None	c.3089C>T	p.P1030L	ENST00000263409	Not reported	29.4	0.647	0.106594602748	Yes
				TNC	None	c.4895G>C	p.G1632A	ENST00000340094	Not reported	29.1	0.6115	0.0378341256418	No
563	BC,39	Yes	No	*RALGDS*	None	c.1184C>T	p.P395L	ENST00000372047	Not reported	25.1	0.602	0.0481839689411	No
				VTI1A	None	c.532T>C	p.S178P	ENST00000393077	Not reported	29.4	0.809	0.129155703701	No
565	OC,43	No	Yes	*EYA2*	rs780344200	c.751G>T	p.G251C	ENST00000317304	Not reported	33.0	0.667	0.152796305407	No
				LPP	rs9830664	c.1717G>A	p.G573S	ENST00000312675	Not reported	33.0	0.694	0.0363434344858	No
				RAD54L	rs369552170	c.1094G>A	p.R365Q	ENST00000371975	Not reported	24.5	0.623	0.0732607632473	No
				ROBO1	rs200951485	c.862T>C	p.S288P	ENST00000436010	Not reported	28.0	0.523	0.0439448821712	No
				WDR11	rs774442833	c.797G>C	p.R266P	ENST00000263461	Not reported	34.0	0.898	0.360584962861	No
581	BC, 46	Yes	No	*FEN1*	rs750703675	c.311G>A	p.R104Q	ENST00000305885	Not reported	30.0	0.587	0.139779212876	No
				NTHL1	rs1805378	c.527T>C	p.I176T	ENST00000219066	Not reported	25.1	0.876	0.51154662686	No
				RGPD3	rs199757971	c.236C>T	p.A79V	ENST00000304514	Not reported	24.6	0.634	0.0812635257962	No
593	BC, 38	Yes	Yes	*BCL2L1*	None	c.485T>C	p.L162S	ENST00000307677	Not reported	26.1	0.624	0.0300581475908	No
626	BC, 46	Yes	Yes	*FOXD1*	None	c.415C>G	p.L139V	ENST00000499003	Not reported	23.4	Not reported	0.865648228319	No
				TLX3	rs200588592	c.695A>G	p.E232G	ENST00000296921	Not reported	24.2	0.552	0.937933441676	No
633	BBC, 38	Yes	No	*FAT4*	None	c.7022G>A	p.C2341Y	ENST00000335110	Not reported	32.0	0.817	0.35606994393	Yes
				MSH6	rs778287080	c.2885T>C	p.I962T	ENST00000234420	Not reported	25.2	0.521	0.15803603547	No
				PAX7	rs201602654	c.602A>G	p.E201G	ENST00000375375	Not reported	24.2	0.531	0.19890393317	Yes
				PTPN13	rs61757790	c.1795G>C	p.D599H	ENST00000316707	Not reported	24.0	0.71	0.0777206298342	Yes
638	BC,42	Yes	No	*EXO1*	rs149397534	c.820G>A	p.G274R	ENST00000348581	Not reported	34.0	0.541	0.0598680279414	No
				NCKIPSD	rs75577765	c.1651C>T	p.P551S	ENST00000294129	Not reported	26.5	0.511	Not reported	No

649	BC, 38	Yes	Yes	*PMS1*	rs1145232	c.856G>A	p.G286R	ENST00000409593	Not-provided, Likely-benign	25.9	0.779	Not reported	No
656	OC, 41 / BC: 48	No	Yes	*COL1A1*	rs760258050	c.2069C>T	p.P690L	ENST00000225964	Not reported	25.6	0.673	0.228892463414	No
				ECT2L	rs144150484	c.2191C>T	p.H731Y	ENST00000367682	Not-provided	25.3	0.517	Not reported	No
				ESCO2	rs115144373	c.1735C>A	p.P579T	ENST00000305188	Uncertain	29.8	0.616	Not reported	No
				IRF4	None	c.338G>A	p.R113Q	ENST00000380956	Not reported	35.0	0.778	0.230099402249	No
				VWA2	None	c.995G>A	p.C332Y	ENST00000392982	Not reported	29.0	0.772	0.152950718841	No
689	BBC, 47	Yes	No	*PLCG1*	rs761555687	c.3446T>C	p.I1149T	ENST00000244007	Not reported	23.6	0.597	0.0996991785708	No
695	OC, 21	Yes	No	*ACACA*	rs543034212	c.541C>T	p.P181S	ENST00000335166	Not reported	29.8	0.756	0.085850935305	No
				CHEK2	rs77130927	c.538C>T	p.R180C	ENST00000328354	Uncertain, Likely-benign, Benign	23.2	0.703	0.0334439898104	Yes
				CYP2C8	rs143386810	c.1150G>A	p.G384S	ENST00000371270	Not reported	26.1	0.741	Not reported	No
				ID3	rs146163818	c.145T>A	p.S49T	ENST00000374561	Not reported	24.1	0.807	0.215403244534	No
				TG	rs142124591	c.3149G>T	p.W1050L	ENST00000220616	Not reported	26.4	0.912	0.157002005399	No
960	BBC, 59 and 70	Yes	No	*EP300*	rs763860567	c.4532A>G	p.N1511S	ENST00000263253	Not reported	27.6	0.72	0.334989871844	Yes
				FAT1	rs146471129	c.9583T>A	p.Y3195N	ENST00000441802	Not reported	28.6	0.805	0.136488920555	Yes
				MYH11	rs137934837	c.4604G>A	p.R1535Q	ENST00000300036	Uncertain	35.0	0.773	0.468129913039	No
				NFKB2	None	c.2531T>C	p.V844A	ENST00000189444	Not reported	27.8	0.743	0.322631496399	No
				TSHR	None	c.1670T>G	p.L557W	ENST00000298171	Not reported	24.5	0.543	0.047309249775	No
974	BC, 46	Yes	No	*EED*	rs772993565	c.154C>T	p.R52C	ENST00000263360	Not reported	29.7	0.598	0.217721061419	No
				MPP3	rs189143886	c.617C>A	p.S206Y	ENST00000398389	Not reported	28.9	0.575	0.0713023116509	No
				RECQL	rs150306543	c.401C>T	p.T134I	ENST00000314748	Not reported	31.0	0.68	0.118377409892	No
				SMAD7	rs773511006	c.713C>T	p.T238M	ENST00000262158	Not reported	25.2	0.735	0.378526534647	No
1014	BC, 42	Yes	No	*KDM5A*	rs201675393	c.1136A>G	p.N379S	ENST00000382815	Not reported	23.3	0.523	0.0873954737136	No
				MYH11	rs137934837	c.4604G>A	p.R1535Q	ENST00000300036	Uncertain	35.0	0.773	0.468129913039	No
				NOTCH2	rs782662504	c.455C>T	p.P152L	ENST00000256646	Not reported	28.5	0.783	0.0592907815833	No
1024	BC, 48	Yes	Yes	*FKBP9*	rs150348129	c.940G>A	p.V314M	ENST00000242209	Not reported	31.0	0.736	Not reported	No
				TYRO3	rs751069987	c.1628C>T	p.S543F	ENST00000263798	Not reported	33.0	0.75	0.161623849858	No

1046	BC, 37	Yes	No	*NCAM1*	rs575893571	c.25G>A	p.A9T	ENST00000524665	Not reported	23.4	Not reported	Not reported	No
				PDGFB	None	c.540T>G	p.C180W	ENST00000331163	Not reported	26.5	0.845	0.692068651029	No
				RBL1	None	c.3050T>C	p.I1017T	ENST00000373664	Not reported	28.6	0.688	0.0504149591879	No
1055	OC, 57	Yes	No	*MSH5*	rs28399977	c.1129C>T	p.L377F	ENST00000375703	Not reported	32.0	0.535	Not reported	No
1095	BC, 43	Yes	No	*CUZD1*	rs36212072	c.854T>C	p.I285T	ENST00000368904	Not reported	22.9	0.572	Not reported	No
				EXO1	rs149397534	c.820G>A	p.G274R	ENST00000348581	Not reported	34.0	0.541	0.0598680279414	No
1096	BC, 46	Yes	No	*CUL7*	rs147493246	c.4762C>A	p.L1588I	ENST00000265348	Uncertain, Benign	32.0	0.592	Not reported	No
				IGF1R	rs45445894	c.1162G>A	p.V388M	ENST00000268035	Not reported	31.0	0.671	0.219496891267	No
				MPP3	rs189143886	c.617C>A	p.S206Y	ENST00000398389	Not reported	28.9	0.575	0.0713023116509	No
1097	OC, 46	Yes	No	*CYP1A1*	rs143070677	c.184G>C	p.A62P	ENST00000379727	Not reported	24.4	0.533	0.0848781419973	No
				MAFG	rs773528916	c.355C>T	p.R119W	ENST00000357736	Not reported	27.5	0.632	0.213158327345	No
				NEMF	None	c.161C>A	p.S54Y	ENST00000298310	Not reported	27.4	0.573	0.0632078601516	No
				RAD50	rs200472836	c.353T>C	p.I118T	ENST00000265335	Uncertain	26.7	0.56	0.0330504082758	No
				WIF1	rs751879347	c.608G>A	p.G203E	ENST00000286574	Not reported	29.3	0.818	0.258590922902	Yes
1109	BC, 21	No	Yes	*ABCB1*	rs773597757	c.3412C>T	p.R1138W	ENST00000265724	Not reported	28.6	0.678	0.111867992005	No
1151	BC, 38	Yes	No	*POLN*	rs34554757	c.2133T>G	p.F711L	ENST00000382865	Not reported	27.2	0.791	Not reported	No
1176	BC, 44	No	Yes	*FANCE*	rs371485747	c.329C>T	p.P110L	ENST00000229769	Not reported	27.1	0.596	0.134019183866	No
				LRP12	rs766182789	c.1385G>A	p.R462H	ENST00000276654	Not reported	34.0	0.782	0.221230884228	No
				RBL1	rs149999468	c.1549C>T	p.R517C	ENST00000344359	Not reported	34.0	0.776	0.129201372601	No
1186	OC, 20	No	No	*JAK2*	rs149705816	c.1759C>A	p.H587N	ENST00000381652	Not reported	32.0	0.628	0.226696637565	Yes
				SAV1	None	c.604C>T	p.P202S	ENST00000324679	Not reported	27.1	0.652	0.147403599117	No
1231	OC, 21	Yes	No	*EGFR*	None	c.352G>T	p.A118S	ENST00000275493	Not reported	23.2	0.545	0.0611834143509	Yes
				ERCC6	rs138758064	c.1801G>A	p.G601S	ENST00000355832	Not reported	34.0	0.722	0.308416648536	No
				HEXB	rs121907984	c.1627G>A	p.A543T	ENST00000261416	Benign, Other	34.0	0.6645	0.266179793844	No
				LASP1	rs746452006	c.206C>T	p.P69L	ENST00000318008	Not reported	34.0	0.564	0.18091148265	No
				MPP3	rs189143886	c.617C>A	p.S206Y	ENST00000398389	Not reported	28.9	0.575	0.0713023116509	No
				NOTCH1	None	c.2369C>T	p.T790I	ENST00000277541	Not reported	22.1	0.598	0.177337835004	No
				PCSK7	rs143209024	c.613G>A	p.G205S	ENST00000320934	Not reported	26.4	0.69	0.229574588797	No

1264	BC, 27	Yes	No	*FAN1*	rs148404807	c.149T>G	p.M50R	ENST00000362065	Not reported	28.4	0.761	0.122156593397	No
				FLNA	rs782292372	c.5965C>T	p.H1989Y	ENST00000344736	Not reported	25.2	0.65	0.859427746523	No
1294	BC, 26	Yes	No	*TPM4*	rs373696275	c.59A>T	p.D20V	ENST00000344824	Not reported	32.0	0.692	0.612621125197	No
1308	BC, 45	Yes	No	*CSMD3*	rs377213732	c.4067A>G	p.D1356G	ENST00000297405	Not reported	28.3	0.605	0.115530903775	No
1326	BC,35	Yes	No	*CHEK2*	rs28909982	c.349A>G	p.R117G	ENST00000328354	Pathogenic, Likely-pathogenic	26.8	0.93	0.217611137759	Yes
1327	BC, 37 / Thyroid, 39	Yes	No	*ATM*	rs35963548	c.1595G>A	p.C532Y	ENST00000278616	Uncertain	23.7	0.578	0.0711442282457	Yes
1455	BC,31	Yes	No	*AIP*	rs145047094	c.47G>A	p.R16H	ENST00000279146	Uncertain, Likely-benign	25.0	0.777	0.339391484095	No
1482	BC, 39	Yes	Yes	*GPHN*	None	c.1471A>T	p.I491F	ENST00000305960	Not reported	23.9	0.824	0.145704784224	No

Subtitles: BBC: bilateral breast cancer; BC: breast cancer; OC: ovarian cancer.

The genes most frequently altered were *MPP3* and *PMS1*. A variant in *MPP3* (c.617C>A), not described by ClinVar, was identified in three unrelated patients (IDs: 974, 1096 and 1231), where two had BC at 46 years of age and one had OC at 21 years of age. All patients with this *MPP3* variant reported a family history of BC. Two unique variants were identified in *PMS1* (Table 2), both found in patients who reported a family history of OC.

We identified a total of 21 unique missense variants in genes described as "hallmarks of cancer" by COSMIC database (Figure 3). Among these genes, we can highlight *CHEK2*, *FAT1*, *ATM*, *EGFR*, *ERBB4*, *KIT*, *KRAS* and *TERT*. Some of these genes have already been associated with hereditary/familial tumors, such as the *ATM* and *CHEK2* genes. The variants identified in CHEK2 have conflicting interpretations of pathogenicity at ClinVar. The missense *CHEK2* variant (c.349A>G; ClinVar: pathogenic/ likely-pathogenic) was identified in a patient with BC at 35 years of age (ID= 1326) and a family history of three BC cases. The other variant in *CHEK2* (c.538C>T; ClinVar: VUS/ likely-benign/ benign) was identified in a patient with OC at 21 years of age (ID= 695), who reported a family history of BC and colorectal cancer. Two different variants were identified in *FAT1* (IDs: 494: c.7957G>A and 960: c.9583T>A), where neither have been described by ClinVar. The variant identified in *ATM* (c.1595G>A) having conflicting interpretations of pathogenicity by ClinVar, was identified in a patient with BC at 37 years of age, and, thyroid cancer at 39 years of age (ID=1327). In addition, this patient reported two relatives BC and one with leukemia. The variant identified in the oncogene *EGFR* (c.352G>T), which was not reported by ClinVar, was identified in patient 1231 with OC at 21 years of age who reported ne case of BC in a first-degree relative. In patient 29, variants in the oncogenes *ERBB4* (c.3446G>T) and *KRAS* (c.461A>G) were observed, where neither have been reported by ClinVar. This patient had OC at 42 years of age and BC at 53 years of age. In addition, this patient reported the presence of several tumors in her family history, especially in first- and second-degree relatives, notably BC, OC, uterine and gastric cancers. The variant found in the oncogene KIT (c.50T>C), reported by ClinVar as VUS, was observed in patient 426. This patient had BC at 38 years of age, and reported a family history of BC and colorectal tumors. The variant identified in *TERT* (c.3332C>T), reported by ClinVar as VUS, was identified in patient 275. This patient had OC at 60 years of age and reported a family history of BC and prostate cancer.

Regarding the genes associated with DNA repair, variants in *EXO1*, *FAN1*, *MLH1*, *MC1R*, *RAD54L* and *RAD50* were identified in our cohort. Two patients (ID: 638 and 1095), both with

BC before the age of 45 shared the same variant in *EXO1* (c.820G>A). Two unrelated patients (426 and 1264) carried the same *FAN1* variant (c.149T>G), and this variant has not been described by ClinVar. Additionally, these patients had BC at 38 and 48 years of age, respectively. The variant identified in *MLH1* (c.794G>A), reported by ClinVar as VUS, was observed in patient 85, who had BC at 51 years of age and reported a family history of BC and gastric cancer. The same *MC1R* variant (c.4604G>A) was observed in patients 960 and 1014. This variant was described by ClinVar as having uncertain significance and was associated with cardiovascular disease. Both *MC1R* variant carriers had BC (patient 960 had two primary BC) and reported a family history of BC. Interestingly, two different missense variants in *RAD54L*, neither reported by Clinvar, were identified in two OC patients (patient 320: c.604C>T and patient 565: c.1094G>A). Both carriers reported a family history of OC. Although only one variant was observed in *RAD50*, a gene involved in DNA repair, this result is noteworthy. The variant (c.353T>C) observed in this study in a patient (1097) with OC at 46 years of age, was reported by ClinVar as uncertain significance, with only two submitters. Details about the personal and family history of the patients with variants in these genes, as well as details regarding the type of variant identified are depicted in Table 2.

Discussion

For a great proportion of the non-*BRCA1/BRCA2* patients with familial BC and/or OC the genetic cause associated with their cancer predisposition is unknown. In the current study, we aimed to identify rare variants that would contribute to HBOC susceptibility in 52 Brazilian high-risk HBOC families BRCA1/BRCA2 mutation-negative by WES and bioinformatic analyses approach.

An initial investigation in our cohort was based on the protein function, prioritizing the analysis of cancer-associated genes. In addition, the variant classification by ClinVar was considered. Variants classified as benign/likely-benign and pathogenic by ClinVar were removed because there is evidence to support their reported classification through in vitro assay, disease mechanism, population studies or segregation studies[21].

A total of 53 unique LoF and 128 unique rare missense variants were identified in the present study. Germline variants were identified in 23 genes described as “hallmarks of cancer”, by COSMIC. Of these, four genes were affected with LoF variants, and, 19 genes were

affected by rare missense variants. We emphasize some of these genes due to their genetic function and the possible association with hereditary/familial tumors.

The most frequently mutated gene with LoF variants was *CCND3*, identified in four unrelated BC patients (7.7% of the cohort). CCND3 encodes a protein that functions in the regulation of cyclin-dependent kinases in the cell cycle[20]. An independent study, reported the presence of the same variant (c.379G>T), in germline tissue, in 1.61% (9/557) of the Caucasians OC patients[22]. Germline variants of CCND3 have not been reported in BC, though 15% (n= 3/20) of metastatic metaplastic breast carcinoma, had an amplification involving CCND3.Interestingly, the authors of that study suggest that *CCND3* could be a potential target of therapy (Nutlins)[23].

The nonsense variant (c.1498G>T) identified in *DROSHA*, was not reported by ClinVar or in the literature. DROSHA has been described as crucial in microRNA biogenesis and, more recently, in translational control and in the direct interaction with p53 effectors associated with RNA binding[20]. Additionally, some authors have described its protein interaction with BRCA1 and involvement in the SMAD3/TP53/DHX9 pathway, which promotes miRNA maturation[24,25]. Interestingly, a single nucleotide polymorphism in the DROSHA gene (rs78393591), was reported in women of African ancestry who had BC[24].Moreover, somatic mutation in *DROSHA* has been shown to be high frequent in Wilms tumor[26]. These observations suggest a possible role for DROSHA in cancer etiology.

A previously not unreported splice acceptor variant identified in the hallmark of cancer-associated gene *SLC34A2* was found in a woman with personal and family history of OC. Study performed by Kanchi *et al*. (2014), through the analysis of germline and somatic variants in OC, observed the presence of a germline splice-site variant (c.1458+2T>C) and a missense variant (c.1079C>T)[22] involved in this suppressor tumor gene. However, as these variants were found at a low in lower frequency in OC cases, the authors could not associate the variant with disease development.

Germline variants in the mismatch DNA repair (MMR) system that involve *MLH1*, *MSH2*, *PMS2*, and *MSH6*, have been associated with Lynch syndrome which feature largely feature colorectal and endometrial cancers, and OC[27,28]. However, some researchers have also identified an association of germline pathogenic variants in MMR genes with BC, although a discrepancy is observed in the literature. A study conducted by Couch *et al.* (2017), analyzing 65,057 non-*BRCA1/BRCA2* mutation carrier women with BC and referred for hereditary cancer

genetic testing, estimated moderate risks of BC with carrier mutations in *MLH1* (OR: 1.15, 95% CI: 0.30-4.19), *MSH2* (OR: 2.46, 95% CI: 0.81-6.93), *MSH6* (OR: 1.93, 95% CI: 1.16-3.27) or *PMS2* (OR: 0.82,95% CI: 0.44-1.47)[29]. In the current study, variants involved in genes *MLH1, MSH2, MSH6* (missense variants) and *PMS2* (frameshift) were also observed in patients with personal and family history of BC. Although the majority of the variants identified in these genes in our study were described by ClinVar as VUS, some were considered pathogenic by submitters this database. Further study is warranted to determine the potential pathogenicity and penetrance of VUS identified in MMR genes for BC risk.

In this study we identified another gene with LoF and previously associated with colorectal cancer in *POLQ* (c.4262_4268del). Interestingly, this was found in two unrelated patients with BC. Studies published by Wang *et al.* (2008)[30] and Brandalize *et al.* (2014)[31] reported the presence of germline variants in *POLQ*, involved patients non-BRCA1/BRCA2 mutated BC cases with a family history of BC. In addition, a case-control study published by Family *et al.* (2015)[32], associated three missense variants in *POLQ* with an increased risk in developing BC. Thus, the authors concluded that *POLQ* could be considered a possible gene candidate involved in the development of hereditary BC.

Other genes related to hallmarks of cancer have also been associated with an increased risk for the development of BC and OC. *ATM* and *CHEK2* are good examples. The rare missense variants identified in these genes in our cohort were described with conflicting interpretations by ClinVar. In 2017, Couch *et al.* showed that pathogenic variants in *ATM* and *CHEK2* were associated with moderate risk of BC (OR: 2.78, 95%CI: 2.22-3.62 and 2.26, 95%CI: 1.89-2.72, respectively)[29]. Similar results were published by Lu *et al.* (2018). The risk to develop BC carriers of mutations in *ATM* was 2.97 (95%CI: 1.67-5.68) and in *CHEK2* was 2.19 (95%CI: 1.40-3.56). Although the authors of that study also estimated the risks of OC, only *ATM* mutation carriers showed an increased risk (OR: 2.85 - 95%CI: 1.30-6.32) for this cancer [33]. An independent study reported that the risk of OC in *CHEK2* mutation carriers was also moderate (OR: 0.86, 95% CI: 0.56-1.33)[34].

We identified variants in *FAT1*, a tumor suppressor gene involved in the WNT/β-catenin pathway[20], in two families with either BC or OC cases. Variants in *FAT1* was also identified in the germline in WES study of 147 cases of colorectal cancer (CRC) from Canadian and Australian families[35]. All patients in that study were diagnosed under the age of 50 years, and, showed the criteria of Familial Colorectal Cancer Type X (FCCTX)[35]. However,

none of our patients with *FAT1* variants reported the presence of colorectal cancer in their family history.

It is important to emphasize that some rare missense variants identified in our study were found in oncogenes, such as *EGFR*, *ERBB4*, *KIT* and *KRAS*, though many of these variants have not been described by ClinVar. These genes are described as kinase superfamily[20]. Somatic amplification and variants in these genes have been molecularly characterized and are established drivers in many cancer types, especially in BC, lung and gastrointestinal stromal cancers, and glioblastomas. However, some reports have described the presence of germline variants in these genes in familial BC/OC. Additionally, Penkert *et al.* (2018), performed next-generation sequencing of 94 cancer predisposition genes in German cohort of 83 unrelated women diagnosed with BC, without germline variants in *BRCA1, BRCA2* and *TP53*, and reported germline missense variants in *EGFR* (c.2039G>A) and KIT (c.391G>A), both with frequency of 1.2%[36]. In our cohort, variants in these genes were found in patients with OC (*EGFR*), both OC and BC cancer (*ERBB4* and *KRAS*) and BC (*KIT*). The variants identified in *KRAS* and *EGFR* were located in domains predicted to activate these proto-oncogenes (Ras and Receptor-L domain, respectively)[37]. The association of the variants identified in these oncogenes in our cohort will be further evaluated by co-segregation analysis.

Some variants involving the *RAD*-family were observed in the current study, were mainly found in probands with personal and family history of OC. The *RAD*-family is involved in the homologous recombination and DNA repair processes[20].Variants in this gene family included *RAD50* (missense), *RAD54L* (missense), *RAD51C* (frameshift), and *RAD54B* (frameshift) where most have not been reported by ClinVar. Some authors have studied the presence of germline variants in these genes in non-*BRCA1/BRCA2* mutation carrier patients at-high risk for BC/OC. Kanchi *et al.*, (2014) described the presence of germline pathogenic variants involving RAD51C and RAD54L in OC cases (both with a frequency of 0.23%)[22]. A targeted mutation screen of *RAD51B*, *RAD51C* and *RAD51D* in 3,429 patients with invasive epithelial OC and 2,000 unaffected *BRCA1/BRCA2* mutation negative women, found rare variants in OC cases in *RAD51C* (14, 0.41%), *RAD51D* (12, 0.35%) and in *RAD51B* (two, 0.06%)[29]. In addition, *RAD51C* variants were associated with a risk of 5.2 (OR, 95% CI: 1.1-24.0)[38]. An increased risk for OC was purported for RAD51C variant carriers (OR, not estimated) in a study by Lu *et al.* (2018)[33]. In contrast Couch *et al*., (2017), showed that potentially pathogenic variants in RAD51C were not associated with increased risk for BC (OR:

0.78, 95% CI: 0.47-1.37)[29]. A recent study by Fan *et al.* (2018) investigating the clinical impact of RAD50 germline variants in 7,657 unselected BC non-*BRCA1/BRCA2* carrier patients reported a carrier frequency of 26/7,657 (0.34%) in patients. Interestingly, the same study reported a multivariate analysis revealing that potentially pathogenic *RAD50* variants were an independent unfavorable predictor of recurrence-free survival (OR: 2.66; 95%CI: 1.18-5.98) and disease-specific survival (OR: 4.36; 95%CI: 1.58-12.03)[39]. Thus, the impacts of the variants in *RAD50* and *RAD51C* on BC susceptibility are still controversial. In our study, the *RAD54B* variant was identified only in a BC case, and the carrier of a variant *RAD51C* developed two primary tumors of the ovary and breast, though this might reflect the rarity of germline carriers of these genes and sample size of cases investigated by WES in our study.

Although WES has become an effective approach to identify variants in alleles and thereby new candidate genetic predisposition genes in many hereditary diseases, it comes with numerous challenges. The different lists of genes that resulted from different breast/ovarian tumors WES studies may be elucidated in part by the different pipelines and bioinformatics tools used to evaluate these data, as well as can be a consequence of the population-specific variability or even due to the selection criteria adopted. Cohorts used to for WES study also play a role in comparing results from different studies. Most notably, we have compared our findings with the Kanchi *et al.* study[22] who reported germline (and somatic) variants in unselected OC cases in contrast to our study of high-risk OC cases from cancer families. Furthermore, WES data users apply different filters to help prioritize variants, thus some variants can be included/removed erroneously.

CONCLUSION

In summary, the present study performed a characterization of germline variants identified in cancer-associated genes, using WES and bioinformatic analyses in Brazilian non-*BRCA1/BRCA2/TP53* mutation-carrier women with BC and/or OC. Our findings suggest that several novel cancer-associated genes also may have a role in HBOC. In addition, the present study provides additional evidence for the association of moderate-risk genes, such as *CHEK2, RAD50, RAD51C,* and *RAD54L*, to the development of familial BC/OC. The candidate genes identified in our high-risk cancer cases would require further validation in Brazilian and other cancer cohorts. Specific variants would also require in vitro analysis to investigate functional consequences on protein function. Such advances will help with the molecular cataloguing of

breast/ovarian tumors in non-*BRCA1/BRCA2* mutation carrier patients as well as the development of gene-based early detection strategies and targeted therapies as has been established in *BRCA1/BRCA2* mutation carrier women.

Acknowledgements

We would like to thank all our patients and their families who contributed to this study. The authors are very grateful to Suzanna L. Arcand, Wejdan Alenezi and Bruna Durães de Figueiredo Barros for whole-exome sequencing analysis.

SUPPLEMENTARY TABLES

Table S1 – Cancer gene reference lists

Gene Symbol	Hallmark	Gene Symbol	Hallmark	Gene Symbol	Hallmark	Gene Symbol	Hallmark
1433S	No	*EIF4A2*	Yes	*MCF2L*	No	*RL40*	No
2AAB	No	*EKI2*	No	*MCL1*	No	*RL7A*	No
A1BG	No	*ELANE*	No	*MCM4*	No	*RLF*	No
A1CF	No	*ELF3*	No	*MCTS1*	No	*RMI1*	No
AAPK1	No	*ELF4*	Yes	*MD1L1*	No	*RMI2*	No
ABCB1	No	*ELF5*	No	*MDC1*	No	*RMP*	No
ABCG2	No	*ELK1*	No	*MDGA1*	No	*RN135*	No
ABI1	Yes	*ELK4*	Yes	*MDM2*	No	*RN139*	No
ABI2	No	*ELL*	Yes	*MDM4*	No	*RN149*	No
ABL1	Yes	*ELN*	No	*MDS2*	No	*RN167*	No
ABL2	No	*ELOC*	No	*MECOM*	No	*RN213*	No
ABLM3	No	*EME1*	No	*MED1*	No	*RN5A*	No
ABRX1	No	*EME2*	No	*MED10*	No	*RNF11*	No
ABTB1	No	*EML4*	Yes	*MED12*	Yes	*RNF168*	No
ACACA	No	*EMP2*	No	*MED28*	No	*RNF213*	Yes
ACBD4	No	*EN113*	No	*MELK*	No	*RNF34*	No
ACK1	No	*ENDOV*	No	*MEN1*	No	*RNF4*	No
ACKR3	Yes	*ENG*	No	*MEP50*	No	*RNF43*	Yes
ACL6A	No	*ENK18*	No	*MERL*	No	*RNF6*	No
ACPH	No	*ENK19*	No	*MERTK*	No	*RNF8*	No
ACS2B	No	*ENK21*	No	*MET*	No	*RNT2*	No
ACSL3	Yes	*ENK24*	No	*MFHA1*	No	*RNZ2*	No
ACSL5	No	*ENK5*	No	*MFR1L*	No	*ROBO1*	No
ACSL6	No	*ENK6*	No	*MGMT*	No	*ROBO2*	No
ACTB	No	*ENK7*	No	*MGP*	No	*RON*	No
ACTN4	No	*ENK8*	No	*MIB2*	No	*ROR1*	No
ACTZ	No	*ENK9*	No	*MIEN1*	No	*RORA*	No
ACV1B	No	*ENL*	No	*MINK1*	No	*ROS1*	Yes
ACVL1	No	*ENO2*	No	*MINP1*	No	*RPA1*	No

ACVR1	Yes	*ENTP5*	No	*MINY3*	No	*RPA2*	No
ACVR2A	No	*EP300*	Yes	*MITF*	No	*RPA3*	No
ADAM17	No	*EPAS1*	Yes	*MK01*	No	*RPA4*	No
ADAP1	No	*EPC1*	No	*MK03*	No	*RPB1*	No
AEN	No	*EPCAM*	No	*MK06*	No	*RPL10*	No
AF10	No	*EPHA2*	No	*MK07*	No	*RPL22*	No
AF17	No	*EPHA3*	No	*MK09*	No	*RPL5*	No
AF1Q	No	*EPHA4*	No	*MK15*	No	*RPN1*	No
AF9	No	*EPHA5*	No	*MKL1*	No	*RRAS2*	No
AFAD	No	*EPHA7*	No	*MKL2*	No	*RRM2B*	No
AFF1	Yes	*EPHA8*	No	*MKRN2*	No	*RRP1B*	No
AFF3	Yes	*EPHB2*	No	*MLF1*	No	*RS27A*	No
AFF4	Yes	*EPHB3*	No	*MLH1*	No	*RS30*	No
AFP	No	*EPHB6*	No	*MLH3*	No	*RSLAA*	No
AGAP2	No	*EPN1*	No	*MLLT1*	No	*RSPO2*	No
AGR2	No	*EPS15*	Yes	*MLLT10*	No	*RSPO3*	No
AGRA2	No	*EPS8*	No	*MLLT11*	No	*RSSA*	No
AGRA3	No	*ERBB2*	Yes	*MLLT3*	No	*RT11*	No
AGRB1	No	*ERBB3*	Yes	*MLLT4*	No	*RTEL1*	No
AGRF1	No	*ERBB4*	Yes	*MLLT6*	No	*RTKN*	No
AGRF2	No	*ERC1*	Yes	*MLX*	No	*RTN4*	No
AGRL2	No	*ERCC1*	No	*MME*	No	*RUBCL*	No
AHRR	No	*ERCC2*	Yes	*MMP1*	No	*RUFY3*	No
AIFM2	No	*ERCC3*	Yes	*MMP14*	No	*RUNDC2A*	No
AIM2	No	*ERCC4*	Yes	*MMP2*	No	*RUNX1*	No
AIP	No	*ERCC5*	Yes	*MMP9*	No	*RUNX1T1*	No
AKAP9	No	*ERCC6*	No	*MMRN2*	No	*RUNX2*	No
AKIP1	No	*ERCC6L*	No	*MMS19*	No	*RUNX3*	No
AKP13	No	*ERCC8*	No	*MN1*	No	*RUVB1*	No
AKT1	Yes	*ERD21*	No	*MNAT1*	No	*RUVB2*	No
AKT2	No	*EREG*	No	*MNX1*	No	*RUXG*	No
AKT3	No	*ERF*	No	*MO4L1*	No	*RXFP2*	No
ALB	No	*ERG*	Yes	*MO4L2*	No	*RYK*	No
ALDH2	No	*ERRFI*	No	*MOB1A*	No	*S100A7*	No
ALEX	No	*ESCO2*	No	*MOD5*	No	*S10A2*	No
ALK	Yes	*ESR1*	Yes	*MOS*	No	*S22AA*	No
ALKBH2	No	*ESR2*	No	*MOV10*	No	*S22AI*	No
ALKBH3	No	*ESX1*	No	*MP2K1*	No	*S2533*	No
ALPK1	No	*ETNK1*	Yes	*MP2K4*	No	*S2547*	No
AMER1	Yes	*ETS1*	No	*MP2K6*	No	*S35B2*	No
AMGO2	No	*ETS2*	No	*MP2K7*	No	*S38A3*	No
AN32A	No	*ETV1*	No	*MPG*	No	*SALL4*	No
ANCHR	No	*ETV3*	No	*MPIP2*	No	*SAMN1*	No
ANDR	No	*ETV4*	No	*MPIP3*	No	*SAPC2*	No
ANK1	No	*ETV5*	No	*MPL*	No	*SASH1*	No
ANM3	No	*ETV6*	Yes	*MPP3*	No	*SAV1*	No
ANM6	No	*ETV7*	No	*MRE11*	No	*SBDS*	No
ANO1	No	*EVI2A*	No	*MRE11A*	No	*SC23B*	No

ANS1A	No	*EVI2B*	No	*MRGBP*	No	*SC31A*	No
ANS1B	No	*EVI5*	No	*MRVI1*	No	*SC5A8*	No
ANXA5	No	*EWS*	No	*MSD3*	No	*SCRIB*	No
APC	Yes	*EWSR1*	Yes	*MSH2*	No	*SCUB3*	No
APC10	No	*EXO1*	No	*MSH3*	No	*SDC1*	No
APCL	No	*EXT1*	No	*MSH4*	No	*SDC4*	No
APEX1	No	*EXT2*	No	*MSH5*	No	*SDCB1*	No
APEX2	No	*EYA2*	No	*MSH6*	No	*SDCG3*	No
APITD1	No	*EZH2*	No	*MSI2*	No	*SDHA*	Yes
APLF	No	*EZR*	Yes	*MSMB*	No	*SDHAF2*	No
APOBEC3B	Yes	*F107A*	No	*MSN*	No	*SDHB*	No
APTX	No	*F10A5*	No	*MSRE*	No	*SDHC*	No
AR	Yes	*F120A*	No	*MTA1*	No	*SDHD*	No
ARAF	No	*F189B*	No	*MTA2*	No	*SE6L1*	No
ARF	No	*F210B*	No	*MTA3*	No	*SEM3B*	No
ARG39	No	*FA32A*	No	*MTCP1*	No	*SEM3F*	No
ARHG1	No	*FA58B*	No	*MTG16*	No	*SEM5A*	No
ARHG4	No	*FA72A*	No	*MTG8*	No	*SENP1*	No
ARHG5	No	*FA83A*	No	*MTG8R*	No	*SENP6*	No
ARHG8	No	*FA83B*	No	*MTND*	No	*SEPT5*	No
ARHGAP26	No	*FA83D*	No	*MTOR*	Yes	*SEPT6*	No
ARHGAP5	No	*FABP7*	No	*MTSS1*	No	*SEPT9*	No
ARHGC	No	*FACD2*	No	*MTUS1*	No	*SESN1*	No
ARHGEF10	No	*FADD*	No	*MUC1*	No	*SESN2*	No
ARHGEF10L	No	*FAK1*	No	*MUC16*	No	*SET*	No
ARHGEF12	No	*FAM131B*	No	*MUC4*	No	*SETBP*	No
ARHGI	No	*FAM135B*	No	*MUL1*	No	*SETBP1*	No
ARHGQ	No	*FAM175A*	No	*MUS81*	No	*SETD1B*	No
ARI3A	No	*FAM3C*	No	*MUTYH*	No	*SETD2*	No
ARI3B	No	*FAM46C*	No	*MXI1*	No	*SETMAR*	No
ARI4A	No	*FAM47C*	No	*MXRA5*	No	*SETMR*	No
ARI5B	No	*FAN1*	No	*MY18B*	No	*SF3B1*	No
ARID1A	Yes	*FANCA*	No	*MYB*	No	*SFPQ*	Yes
ARID1B	No	*FANCB*	No	*MYBA*	No	*SFRP4*	No
ARID2	Yes	*FANCC*	No	*MYBB*	No	*SG2A2*	No
ARK73	No	*FANCD2*	Yes	*MYC*	Yes	*SGK1*	No
ARK74	No	*FANCE*	No	*MYCL*	No	*SH21B*	No
ARNT	Yes	*FANCF*	No	*MYCN*	No	*SH2B1*	No
ASC	No	*FANCG*	No	*MYCP1*	No	*SH2B2*	No
ASF1A	No	*FANCI*	No	*MYD88*	No	*SH2B3*	No
ASPC1	No	*FANCJ*	No	*MYEOV*	No	*SH3G1*	No
ASPP2	No	*FANCL*	No	*MYH11*	No	*SH3GL1*	No
ASPSCR1	Yes	*FANCM*	No	*MYH9*	No	*SH3R2*	No
ASXL1	Yes	*FAS*	Yes	*MYO5A*	No	*SHB*	No
ASXL2	No	*FASLG*	No	*MYOD1*	No	*SHC1*	No
ATAD2	No	*FAT1*	Yes	*N4BP2*	No	*SHC2*	No
ATF1	Yes	*FAT3*	No	*NAA15*	No	*SHC3*	No
ATF5	No	*FAT4*	Yes	*NAB2*	Yes	*SHCAF*	No

ATF7	No	*FBLN1*	No	*NACA*	No	*SHFM1*	No
ATIC	Yes	*FBLN2*	No	*NACC2*	No	*SHIP2*	No
ATM	Yes	*FBSP1*	No	*NADAP*	No	*SHPRH*	No
ATP1A1	Yes	*FBW1A*	No	*NANO1*	No	*SHSA5*	No
ATP2B3	Yes	*FBW1B*	No	*NANP8*	No	*SIAH1*	No
ATR	Yes	*FBX5*	No	*NARR*	No	*SIK1*	No
ATRIP	No	*FBX7*	No	*NAT6*	No	*SIK3*	No
ATRX	Yes	*FBXL2*	No	*NAV2*	No	*SIR1*	No
ATXN3	No	*FBXO11*	No	*NBAS*	No	*SIR2*	No
AURKA	No	*FBXW7*	Yes	*NBEA*	No	*SIR4*	No
AURKB	No	*FCG2B*	No	*NBL1*	No	*SIRPA*	No
AXIN1	Yes	*FCGR2B*	Yes	*NBN*	No	*SIVA*	No
AXIN2	Yes	*FCRL4*	No	*NBPF3*	No	*SIX1*	No
B2CL2	No	*FCSD2*	No	*NBPFC*	No	*SIX2*	No
B2L10	No	*FEM1B*	No	*NCAM1*	No	*SKI*	No
B2LA1	No	*FEN1*	No	*NCKIPSD*	No	*SKIL*	No
B2M	Yes	*FER*	No	*NCOA1*	No	*SLAP2*	No
BACH2	No	*FES*	No	*NCOA2*	No	*SLC34A2*	Yes
BAD	No	*FEV*	No	*NCOA4*	No	*SLC45A3*	Yes
BAG3	No	*FGF1*	No	*NCOR1*	No	*SLIP*	No
BAK	No	*FGF10*	No	*NCOR2*	Yes	*SLX1A*	No
BANP	No	*FGF2*	No	*NDC80*	No	*SLX1B*	No
BAP1	Yes	*FGF22*	No	*NDE1*	No	*SLX4*	No
BARD1	No	*FGF3*	No	*NDEL1*	No	*SMAD2*	Yes
BATF	No	*FGF4*	No	*NDKA*	No	*SMAD3*	Yes
BATF2	No	*FGF5*	No	*NDKB*	No	*SMAD4*	No
BATF3	No	*FGF6*	No	*NDRG1*	Yes	*SMAD5*	No
BAX	No	*FGF7*	No	*NDRG2*	No	*SMAD6*	No
BAZ1A	No	*FGF8*	No	*NDUAD*	No	*SMAD7*	No
BC11A	No	*FGFP1*	No	*NDUC2*	No	*SMAGP*	No
BC11B	No	*FGFR1*	Yes	*NDUF4*	No	*SMARCA4*	No
BCAR1	No	*FGFR1OP*	No	*NEB2*	No	*SMARCB1*	No
BCAS3	No	*FGFR2*	Yes	*NEDD4*	No	*SMARCD1*	No
BCAS4	No	*FGFR3*	Yes	*NEDD8*	No	*SMARCE1*	No
BCCIP	No	*FGFR4*	Yes	*NEIL1*	No	*SMC1A*	No
BCL10	Yes	*FGOP2*	No	*NEIL2*	No	*SMC3*	No
BCL11A	Yes	*FGR*	No	*NEIL3*	No	*SMC5*	No
BCL11B	Yes	*FH*	No	*NEK2*	No	*SMCA4*	No
BCL2	No	*FHIT*	Yes	*NEK3*	No	*SMO*	No
BCL2L1	No	*FHL2*	No	*NEK4*	No	*SMUF2*	No
BCL2L12	No	*FIP1L1*	Yes	*NEMF*	No	*SMUG1*	No
BCL3	No	*FKBP9*	No	*NEMO*	No	*SNAI1*	No
BCL6	No	*FLCN*	No	*NENF*	No	*SNAI2*	No
BCL6B	No	*FLI1*	No	*NEO1*	No	*SND1*	No
BCL7A	No	*FLNA*	No	*NEUL1*	No	*SNF5*	No
BCL9	Yes	*FLOT1*	No	*NF1*	Yes	*SNIP1*	No
BCL9L	No	*FLT1*	No	*NF2*	Yes	*SNTB1*	No
BCLAF1	No	*FLT3*	Yes	*NFAT5*	No	*SNW1*	No

BCOR	Yes	*FLT3L*	No	*NFATC2*	No	*SOCS1*	No
BCORL1	Yes	*FLT4*	Yes	*NFE2L2*	Yes	*SOX2*	No
BCR	No	*FNBP1*	No	*NFIB*	No	*SOX21*	No
BECN1	No	*FNIP1*	No	*NFIP1*	No	*SP1*	No
BIK	No	*FNIP2*	No	*NFIP2*	No	*SP100*	No
BIN1	No	*FOLH1*	No	*NFKB1*	No	*SPAG1*	No
BIRC3	Yes	*FOS*	No	*NFKB2*	No	*SPECC1*	No
BIRC6	No	*FOSL1*	No	*NFKBIE*	No	*SPEN*	No
BLCAP	No	*FOSL2*	No	*NGAL*	No	*SPI1*	No
BLK	No	*FOXA1*	Yes	*NGF*	No	*SPN90*	No
BLM	Yes	*FOXD1*	No	*NHEJ1*	No	*SPO11*	No
BLNK	No	*FOXE1*	No	*NIN*	No	*SPOP*	Yes
BMI1	No	*FOXL2*	No	*NINL*	No	*SPRTN*	No
BMP5	No	*FOXO1*	No	*NKX21*	No	*SPT13*	No
BMPR1A	Yes	*FOXO3*	No	*NKX2-1*	No	*SQSTM*	No
BMR1A	No	*FOXO4*	No	*NKX31*	No	*SRC*	No
BMR1B	No	*FOXP1*	No	*NKX3-1*	No	*SRC8*	No
BMX	No	*FOXR1*	No	*NLS1*	No	*SRGAP3*	No
BOREA	No	*FR1OP*	No	*NMI*	No	*SRGP1*	No
BRAF	Yes	*FRAT1*	No	*NOL11*	No	*SRSF1*	No
BRCA1	Yes	*FRK*	No	*NOL7*	No	*SRSF2*	No
BRCA2	Yes	*FRMD3*	No	*NONO*	Yes	*SRSF3*	No
BRCC3	No	*FSCN1*	No	*NOP53*	No	*SRY*	No
BRD1	No	*FSTL3*	No	*NOTC4*	No	*SS18*	No
BRD3	No	*FUBP1*	Yes	*NOTCH1*	No	*SS18L1*	No
BRD4	Yes	*FUMH*	No	*NOTCH2*	No	*SSPN*	No
BRD7	No	*FUND2*	No	*NOV*	No	*SSX1*	No
BRD8	No	*FURIN*	No	*NPAT*	No	*SSX2*	No
BRI3B	No	*FUS*	No	*NPM*	No	*SSX4*	No
BRIP1	Yes	*FYN*	No	*NPM1*	No	*SSXT*	No
BRMS1	No	*FZD1*	No	*NPRL2*	No	*ST134*	No
BRNP1	No	*FZR1*	No	*NR4A3*	No	*ST14*	No
BTBD12	No	*G3BP1*	No	*NRAS*	Yes	*ST17A*	No
BTC	No	*GA45G*	No	*NRG1*	No	*ST18*	No
BTG1	No	*GADD45A*	No	*NS1BP*	No	*ST20*	No
BTG2	No	*GALNT12*	No	*NSA2*	No	*ST5*	No
BTG3	No	*GAS7*	Yes	*NSD1*	No	*ST7*	No
BTK	Yes	*GATA1*	No	*NSD2*	No	*ST7L*	No
BUB1B	Yes	*GATA2*	No	*NSD3*	No	*STA13*	No
C15orf65	No	*GATA3*	No	*NT5C2*	No	*STABP*	No
C17orf70	No	*GCNT3*	No	*NTHL1*	No	*STAG1*	No
C19orf40	No	*GCR*	No	*NTRK1*	No	*STAG2*	No
C1orf86	No	*GDS1*	No	*NTRK3*	No	*STAP2*	No
C1QBP	No	*GEN1*	No	*NU214*	No	*STAR8*	No
C1TC	No	*GFI1*	No	*NUAK1*	No	*STAT3*	No
C2D1A	No	*GFI1B*	No	*NUDT1*	No	*STAT5B*	No
C2orf44	No	*GGA1*	No	*NUMA1*	No	*STAT6*	No
C56D2	No	*GIPC1*	No	*NUP214*	No	*STEA3*	No

C7orf11	No	*GIT1*	No	*NUP98*	No	*STEA4*	No
CA052	No	*GLI1*	No	*NUTM1*	No	*STIL*	No
CA053	No	*GLI2*	No	*NUTM2A*	No	*STK11*	No
CA2D2	No	*GLT12*	No	*NUTM2B*	No	*STK25*	No
CA2D3	No	*GML*	No	*OBF1*	No	*STK26*	No
CACNA1D	Yes	*GMPS*	No	*OBFC2B*	No	*STK3*	No
CADH1	No	*GNA11*	Yes	*OBSCN*	No	*STK38*	No
CADH2	No	*GNAQ*	Yes	*OGG1*	No	*STK39*	No
CADH3	No	*GNAS*	Yes	*OGR1*	No	*STK4*	No
CADM1	No	*GNAS2*	No	*OLIG2*	No	*STRN*	No
CADM3	No	*GNL3L*	No	*OMD*	No	*STYK1*	No
CADM4	No	*GO45*	No	*OPCM*	No	*SUFU*	No
CAH9	No	*GOGA5*	No	*OTU7B*	No	*SUMO1*	No
CALCA	No	*GOLGA5*	No	*P2RY8*	No	*SUSD2*	No
CALR	Yes	*GOPC*	No	*P3H1*	No	*SUSD3*	No
CAMTA1	Yes	*GOT1B*	No	*P53*	No	*SUSD6*	No
CANT1	Yes	*GP15L*	No	*P73*	No	*SUV91*	No
CARD11	Yes	*GP5*	No	*P85A*	No	*SUV92*	No
CARL3	No	*GPAT3*	No	*P85B*	No	*SUZ12*	Yes
CARS	Yes	*GPC3*	Yes	*PA216*	No	*SYCC*	No
CASC3	No	*GPC5*	No	*PA2G4*	No	*SYCP3*	No
CASC5	Yes	*GPHN*	No	*PABPC1*	No	*SYK*	No
CASL	No	*GPHRA*	No	*PAF1*	No	*SYNP2*	No
CASP2	No	*GPKOW*	No	*PAF15*	No	*SYP*	No
CASP3	No	*GPS2*	No	*PAFAH1B2*	Yes	*SZRD1*	No
CASP6	No	*GRAP*	No	*PAK1*	No	*T184B*	No
CASP8	Yes	*GRB10*	No	*PAK4*	No	*T53I1*	No
CASP9	No	*GRB2*	No	*PAK5*	No	*T53I2*	No
CASPA	No	*GRB7*	No	*PAL4A*	No	*TACC1*	No
CASR	No	*GREM1*	No	*PALB2*	No	*TACC2*	No
CAV1	No	*GRHL2*	No	*PALLD*	No	*TACC3*	No
CAVN3	No	*GRIN2A*	No	*PANO1*	No	*TAD2B*	No
CBFA2T3	No	*GRM3*	No	*PAR10*	No	*TAF12*	No
CBFB	Yes	*GRP1*	No	*PAR6A*	No	*TAF15*	No
CBL	Yes	*GRWD1*	No	*PARK7*	No	*TAF4*	No
CBLB	Yes	*GSDMA*	No	*PARN*	No	*TAL1*	No
CBLC	No	*GSDME*	No	*PARP1*	No	*TAL2*	No
CBLL2	No	*GSTM1*	No	*PARP2*	No	*TAOK1*	No
CBP	No	*GSTP1*	No	*PARP3*	No	*TAOK2*	No
CBX8	No	*GTF2H1*	No	*PARP4*	No	*TAOK3*	No
CC85B	No	*GTF2H2*	No	*PATZ1*	No	*TARG1*	No
CCAR2	No	*GTF2H3*	No	*PAWR*	No	*TAXB1*	No
CCD26	No	*GTF2H4*	No	*PAX3*	Yes	*TBC3A*	No
CCD34	No	*GTF2H5*	No	*PAX5*	Yes	*TBK1*	No
CCDB1	No	*GUAA*	No	*PAX7*	Yes	*TBL1XR1*	Yes
CCDC6	Yes	*H2AFX*	No	*PAX8*	Yes	*TBP*	No
CCL2	No	*H3F3A*	Yes	*PAXI*	No	*TBRG1*	No
CCNB1	No	*H3F3B*	No	*PB1*	No	*TBX3*	No

CCNB1IP1	Yes	*HABP2*	No	*PBIP1*	No	*TCAL7*	No
CCNC	No	*HASP*	No	*PBRM1*	No	*TCAM1*	No
CCND1	Yes	*HBA1*	No	*PBX1*	Yes	*TCAM2*	No
CCND2	Yes	*HCK*	No	*PBX2*	No	*TCEA1*	No
CCND3	No	*HDAC1*	No	*PBX3*	No	*TCF12*	No
CCNE1	Yes	*HDAC2*	No	*PC11Y*	No	*TCF3*	No
CCNE2	No	*HDGR2*	No	*PCA3*	No	*TCF7L2*	Yes
CCNG1	No	*HEAT6*	No	*PCBP1*	No	*TCHP*	No
CCNG2	No	*HELQ*	No	*PCD15*	No	*TCL1A*	Yes
CCNH	No	*HEMK1*	No	*PCLI1*	No	*TCL1B*	No
CCNL1	No	*HERC1*	No	*PCM1*	No	*TCP1L*	No
CCNT2	No	*HERPUD1*	No	*PCNA*	No	*TCTA*	No
CCR4	No	*HEXB*	No	*PCSK7*	No	*TDG*	No
CCR7	No	*HEXI1*	No	*PCX2*	No	*TDP1*	No
CD19	No	*HEXI2*	No	*PDCD1LG2*	Yes	*TDP2*	No
CD209	No	*HEY1*	Yes	*PDCD4*	No	*TEC*	No
CD274	No	*HGF*	No	*PDE4DIP*	No	*TEF*	No
CD28	No	*HIC1*	No	*PDGFB*	No	*TEFF1*	No
CD34	No	*HIC2*	No	*PDGFC*	No	*TENS4*	No
CD38	No	*HIF1A*	Yes	*PDGFD*	No	*TERC*	No
CD3Z	No	*HIF3A*	No	*PDGFRA*	No	*TERT*	Yes
CD40LG	No	*HIP1*	Yes	*PDGFRB*	No	*TES*	No
CD44	No	*HIPK2*	No	*PDLI5*	No	*TET1*	Yes
CD5	No	*HIST1H3B*	No	*PDPN*	No	*TET2*	No
CD74	No	*HIST1H4I*	No	*PDRG1*	No	*TEX10*	No
CD79A	Yes	*HLA-A*	No	*PDZD4*	No	*TF2H1*	No
CD79B	Yes	*HLF*	No	*PEA15*	No	*TF65*	No
CDC23	No	*HLTF*	No	*PEBB*	No	*TF7L2*	No
CDC37	No	*HMGA1*	No	*PEG10*	No	*TFDP1*	No
CDC42	No	*HMGA2*	No	*PER1*	Yes	*TFDP2*	No
CDC73	Yes	*HMGB1*	No	*PERP*	No	*TFDP3*	No
CDCP1	No	*HMGB2*	No	*PFD3*	No	*TFE2*	No
CDH1	Yes	*HMGN2P46*	No	*PGDH*	No	*TFE3*	Yes
CDH10	No	*HNF1A*	Yes	*PGFRA*	No	*TFEB*	No
CDH11	Yes	*HNF1B*	No	*PGFRB*	No	*TFG*	No
CDH17	No	*HNRNPA2B1*	Yes	*PGFRL*	No	*TFIP8*	No
CDH2	No	*HOOK3*	No	*PGM1*	No	*TFPT*	No
CDK1	No	*HOP*	No	*PGR*	No	*TFRC*	No
CDK10	No	*HOT*	No	*PGRP2*	No	*TG*	No
CDK12	Yes	*HOXA11*	Yes	*PHB*	No	*TGFA*	No
CDK2	No	*HOXA13*	No	*PHF23*	No	*TGFB1*	No
CDK20	No	*HOXA9*	No	*PHF6*	No	*TGFBR2*	Yes
CDK4	Yes	*HOXB13*	No	*PHLA3*	No	*TGFR1*	No
CDK6	Yes	*HOXC11*	No	*PHLP1*	No	*TGFR2*	No
CDK7	No	*HOXC13*	No	*PHLP2*	No	*THA*	No
CDK9	No	*HOXD11*	No	*PHOX2B*	Yes	*THADA*	No
CDKA1	No	*HOXD13*	No	*PIAS1*	No	*THAP1*	No
CDKL1	No	*HPGDS*	No	*PIAS2*	No	*THB*	No

CDKL2	No	*HRAS*	Yes	*PIAS4*	No	*THEM4*	No
CDKN1A	No	*HS90B*	No	*PICAL*	No	*THOC1*	No
CDKN1B	No	*HSP90AA1*	No	*PICALM*	Yes	*THRAP3*	No
CDKN1C	No	*HSP90AB1*	No	*PIDD1*	No	*TIAM1*	No
CDKN2A	Yes	*HTAI2*	No	*PIGU*	No	*TIE2*	No
CDKN2C	No	*HTATIP2*	No	*PIK3CA*	Yes	*TIF1A*	No
CDKN3	No	*HUS1*	No	*PIK3CB*	Yes	*TIFA*	No
CDN1A	No	*HUTI*	No	*PIK3R1*	Yes	*TINF2*	No
CDN1B	No	*HUTU*	No	*PIM1*	Yes	*TIPRL*	No
CDN1C	No	*HXA9*	No	*PIM2*	No	*TISB*	No
CDN2A	No	*HXB13*	No	*PIM3*	No	*TISD*	No
CDN2B	No	*HYAL1*	No	*PIMRE*	No	*TLK1*	No
CDN2C	No	*HYAL3*	No	*PIN1*	No	*TLK2*	No
CDN2D	No	*I17RB*	No	*PININ*	No	*TLX1*	No
CDON	No	*IASPP*	No	*PINK1*	No	*TLX3*	No
CDT1	No	*ID1*	No	*PINX1*	No	*TM101*	No
CDX2	No	*ID3*	No	*PIR*	No	*TM102*	No
CE162	No	*IDH1*	Yes	*PIWL1*	No	*TM115*	No
CEACAM5	No	*IDH2*	Yes	*PIWL2*	No	*TM127*	No
CEAM1	No	*IER2*	No	*PK3CA*	No	*TM158*	No
CEAM5	No	*IF16*	No	*PK3CB*	No	*TM9S4*	No
CEAM6	No	*IF2B3*	No	*PK3CD*	No	*TMED4*	No
CEBPA	No	*IF4E*	No	*PKHG2*	No	*TMED8*	No
CEBPG	No	*IF5A2*	No	*PKHG5*	No	*TMEM127*	No
CENPK	No	*IFFO1*	No	*PKHO1*	No	*TMF1*	No
CENPU	No	*IFIX*	No	*PLAG1*	No	*TMPRSS2*	Yes
CENPW	No	*IFM1*	No	*PLAK*	No	*TNAP3*	No
CEP57	No	*IFM3*	No	*PLAL1*	No	*TNC*	No
CEP89	No	*IFNA1*	No	*PLCE1*	No	*TNF*	No
CETN2	No	*IFNG*	No	*PLCG1*	No	*TNF15*	No
CHAF1A	No	*IFRD2*	No	*PLGF*	No	*TNFAIP3*	Yes
CHAF1B	No	*IGF1*	No	*PLK1*	No	*TNFRSF10B*	No
CHCHD7	No	*IGF1R*	No	*PLK2*	No	*TNFRSF14*	No
CHD1L	No	*IGF2*	No	*PLK3*	No	*TNFRSF17*	No
CHD2	No	*IGF2BP2*	No	*PLPL3*	No	*TNFRSF8*	No
CHD4	Yes	*IGH*	Yes	*PLPP5*	No	*TNIP2*	No
CHD5	No	*IGK*	No	*PLXB1*	No	*TNIP3*	No
CHDH	No	*IGL*	No	*PLXB2*	No	*TNK1*	No
CHEK1	No	*IKBB*	No	*PLXB3*	No	*TNR17*	No
CHEK2	Yes	*IKBKB*	No	*PML*	No	*TNR1A*	No
CHFR	No	*IKZF1*	No	*PMS1*	No	*TNR6A*	No
CHGA	No	*IL10*	No	*PMS2*	Yes	*TNR6B*	No
CHIC2	No	*IL15*	No	*PMS2L3*	No	*TNR6C*	No
CHK1	No	*IL18*	No	*PNKP*	No	*TOB1*	No
CHK2	No	*IL1B*	No	*PO3F2*	No	*TOB2*	No
CHST11	No	*IL2*	No	*PO4F1*	No	*TOP1*	No
CHSTB	No	*IL21R*	No	*PO4F2*	No	*TOP3A*	No
CIB1	No	*IL24*	No	*POLA1*	No	*TOPB1*	No

CIC	Yes	*IL3RB*	No	*POLB*	No	*TOPBP1*	No
CIITA	Yes	*IL4RA*	No	*POLD1*	No	*TOPRS*	No
CIP1	No	*IL6*	No	*POLD2*	No	*TP4A2*	No
CIP2A	No	*IL6ST*	Yes	*POLD3*	No	*TP53*	Yes
CJ090	No	*IL7R*	No	*POLD4*	No	*TP53BP1*	No
CK095	No	*IL8*	No	*POLE*	No	*TP63*	Yes
CKAP2	No	*ILK*	No	*POLE2*	No	*TPD52*	No
CLD7	No	*ILKAP*	No	*POLE3*	No	*TPD53*	No
CLIP1	Yes	*ING1*	No	*POLE4*	No	*TPGS2*	No
CLK2	No	*ING2*	No	*POLG*	No	*TPM3*	No
CLOCK	No	*ING3*	No	*POLH*	No	*TPM4*	No
CLP1	No	*ING4*	No	*POLI*	No	*TPO*	No
CLSPN	No	*INSL3*	No	*POLK*	No	*TPOR*	No
CLTC	Yes	*INSM1*	No	*POLL*	No	*TPR*	No
CLTCL1	Yes	*INSR*	No	*POLM*	No	*TR10B*	No
CLUA1	No	*INT6*	No	*POLN*	No	*TRA*	No
CLUS	No	*IRF1*	No	*POLQ*	No	*TRADD*	No
CMC4	No	*IRF3*	No	*POT1*	No	*TRAF1*	No
CMTA1	No	*IRF4*	No	*POU2AF1*	Yes	*TRAF3*	No
CNBD1	No	*IRF7*	No	*POU5F1*	Yes	*TRAF5*	No
CNBP	Yes	*IRS1*	No	*PP1A*	No	*TRAF6*	No
CNOT3	Yes	*IRS4*	No	*PPARG*	Yes	*TRAF7*	No
CNTNAP2	No	*ISK1*	No	*PPFIBP1*	No	*TRB*	No
CNTRL	No	*IST1*	No	*PPIE*	No	*TRD*	No
COL1A1	No	*ISX*	No	*PPM1D*	Yes	*TREX1*	No
COL2A1	No	*ITA9*	No	*PPP2R1A*	No	*TREX2*	No
COL3A1	No	*ITCH*	No	*PPP6C*	Yes	*TRG-GCC2*	No
COMD1	No	*ITGAV*	No	*PPR18*	No	*TRI13*	No
COP1	No	*ITK*	No	*PRAF3*	No	*TRI22*	No
COX1	No	*ITPA*	No	*PRAM*	No	*TRI27*	No
COX6C	No	*IWS1*	No	*PRC1*	No	*TRI33*	No
CPEB3	No	*JAK1*	Yes	*PRCC*	No	*TRI35*	No
CPEB4	No	*JAK2*	Yes	*PRDM1*	No	*TRI36*	No
CPNE1	No	*JAK3*	Yes	*PRDM16*	Yes	*TRI37*	No
CPNE3	No	*JAZF1*	No	*PRDM2*	No	*TRIB3*	No
CQ080	No	*JIP4*	No	*PRDM5*	No	*TRIM24*	No
CR032	No	*JTB*	No	*PRDX6*	No	*TRIM27*	No
CR3L2	No	*JUN*	No	*PREX2*	Yes	*TRIM33*	No
CR3L3	No	*JUND*	No	*PRF1*	No	*TRIM8*	No
CRADD	No	*JUPI2*	No	*PRKACA*	Yes	*TRIP11*	No
CRBL2	No	*KANK1*	No	*PRKAR1A*	Yes	*TRRAP*	No
CRCM	No	*KAP0*	No	*PRKCB*	No	*TS101*	No
CREB1	No	*KAPCB*	No	*PRKDC*	No	*TSC1*	No
CREB3L1	Yes	*KAT2A*	No	*PRKN*	No	*TSC2*	No
CREB3L2	Yes	*KAT2B*	No	*PRKX*	No	*TSHR*	No
CREB5	No	*KAT5*	No	*PROM1*	No	*TSN31*	No
CREBBP	Yes	*KAT6A*	No	*PRPF19*	No	*TSN6*	No
CREG1	No	*KAT6B*	No	*PRPF40B*	No	*TTC23*	No

CRK	No	*KAT7*	No	*PRR14*	No	*TTK*	No
CRKL	No	*KC1D*	No	*PRR5*	No	*TTP*	No
CRLF2	Yes	*KC1E*	No	*PRRX1*	No	*TUSC1*	No
CRNKL1	No	*KC1G2*	No	*PRUN1*	No	*TUSC2*	No
CRPAK	No	*KCC4*	No	*PRUN2*	No	*TX1B3*	No
CRTC1	Yes	*KCD11*	No	*PSA5*	No	*TXK*	No
CRTC3	No	*KCD21*	No	*PSCA*	No	*TXNIP*	No
CRY1	No	*KCNA1*	No	*PSD10*	No	*TYDP2*	No
CSDE1	No	*KCNJ5*	No	*PSIP1*	Yes	*TYK2*	No
CSF1R	No	*KCTD6*	No	*PSMD6*	No	*TYRO*	No
CSF2	No	*KDM1A*	No	*PTC1*	No	*TYRO3*	No
CSF3	No	*KDM3B*	No	*PTC2*	No	*TYW4*	No
CSF3R	No	*KDM5A*	No	*PTCH1*	No	*TYY1*	No
CSK	No	*KDM5C*	No	*PTEN*	Yes	*TZAP*	No
CSK21	No	*KDM6A*	Yes	*PTGS2*	No	*U2AF1*	No
CSK22	No	*KDR*	Yes	*PTHB1*	No	*U2QL1*	No
CSK23	No	*KDSR*	No	*PTK6*	Yes	*UB2D2*	No
CSMD3	No	*KEAP1*	Yes	*PTK7*	No	*UB2R1*	No
CSN3	No	*KHDR1*	No	*PTN13*	No	*UB2R2*	No
CSN5	No	*KI20B*	No	*PTN14*	No	*UB2V2*	No
CSN6	No	*KIAA1549*	No	*PTN18*	No	*UBB*	No
CSPP1	No	*KIAA1598*	No	*PTN7*	No	*UBC*	No
CSRN1	No	*KIF1B*	No	*PTOV1*	No	*UBC9*	No
CTBP1	No	*KIF22*	No	*PTPN11*	No	*UBD*	No
CTCF	Yes	*KIF5B*	No	*PTPN11*	No	*UBE2A*	No
CTDS2	No	*KILIN*	No	*PTPN13*	Yes	*UBE2B*	No
CTDSL	No	*KIME*	No	*PTPN6*	No	*UBE2N*	No
CTIP	No	*KISS1*	No	*PTPRB*	No	*UBE2T*	No
CTLA4	No	*KIT*	Yes	*PTPRC*	No	*UBE2V1*	No
CTNB1	No	*KLF4*	Yes	*PTPRD*	No	*UBE2V2*	No
CTND1	No	*KLF5*	No	*PTPRE*	No	*UBE4B*	No
CTNNA1	No	*KLF6*	No	*PTPRH*	No	*UBIA1*	No
CTNNA2	No	*KLH20*	No	*PTPRJ*	No	*UBIM*	No
CTNNB1	No	*KLH22*	No	*PTPRK*	No	*UBP10*	No
CTNND1	No	*KLH41*	No	*PTPRN*	No	*UBP2*	No
CTNND2	No	*KLK10*	No	*PTPRO*	No	*UBP28*	No
CUL1	No	*KLK2*	No	*PTPRT*	Yes	*UBP32*	No
CUL2	No	*KLK3*	No	*PTPRU*	No	*UBP4*	No
CUL3	No	*KMT2A*	No	*PTTG1*	No	*UBP47*	No
CUL4A	No	*KMT2B*	No	*PTTG2*	No	*UBP6*	No
CUL4B	No	*KMT2C*	No	*PTTG3*	No	*UBR5*	No
CUL5	No	*KMT2D*	No	*PUM1*	No	*UBS3A*	No
CUL7	No	*KMT2E*	No	*PUM2*	No	*UCHL1*	No
CUX1	Yes	*KMT5A*	No	*PWWP2A*	No	*UFL1*	No
CUZD1	No	*KNL1*	No	*PXMP4*	No	*UFO*	No
CXCL12	No	*KNSTRN*	No	*QKI*	Yes	*UHRF1*	No
CXCR4	Yes	*KPCA*	No	*R144B*	No	*UHRF2*	No
CXXC5	No	*KPCD*	No	*RA51C*	No	*ULA1*	No

CYCS	No	*KPCD2*	No	*RA51D*	No	*UNG*	No
CYLD	No	*KPCI*	No	*RA54B*	No	*URFB1*	No
CYP19A1	No	*KPCL*	No	*RAB11B*	No	*US6NL*	No
CYP1A1	No	*KPYM*	No	*RAB26*	No	*USE1*	No
CYP2C8	No	*KRAS*	Yes	*RAB5A*	No	*USP1*	No
CYR61	No	*KRIT1*	No	*RAB7A*	No	*USP44*	No
CYSLTR2	No	*KRT19*	No	*RAB8A*	No	*USP6*	No
CYTSB	No	*KRT7*	No	*RABEP1*	Yes	*USP8*	Yes
DAB2	No	*KS6A2*	No	*RAC1*	Yes	*UVSSA*	No
DAB2P	No	*KS6A5*	No	*RACK1*	No	*VATH*	No
DACH1	No	*KS6B1*	No	*RAD1*	No	*VAV*	No
DACT1	No	*KS6B2*	No	*RAD17*	No	*VAV1*	No
DAPK2	No	*KSYK*	No	*RAD18*	No	*VAV2*	No
DAPK3	No	*KTN1*	No	*RAD21*	Yes	*VAV3*	No
DAXX	Yes	*L2GL1*	No	*RAD23A*	No	*VEGFA*	No
DCAF1	No	*LACTB*	No	*RAD23B*	No	*VEGFC*	No
DCAF12L2	No	*LAP4B*	No	*RAD50*	No	*VGFR1*	No
DCBD2	No	*LARP4B*	No	*RAD51*	No	*VGFR2*	No
DCC	No	*LASP1*	No	*RAD51B*	No	*VGFR3*	No
DCLRE1A	No	*LATS1*	No	*RAD51C*	No	*VHL*	Yes
DCLRE1B	No	*LATS2*	No	*RAD51D*	No	*VIME*	No
DCLRE1C	No	*LC7L3*	No	*RAD52*	No	*VMA5A*	No
DCNL1	No	*LCK*	Yes	*RAD54*	No	*VMP1*	No
DCNL3	No	*LCP1*	No	*RAD54B*	No	*VOPP1*	No
DCR1B	No	*LDB1*	No	*RAD54L*	No	*VRK1*	No
DCTN1	No	*LEF1*	Yes	*RAD9A*	No	*VTI1A*	No
DDB1	No	*LEG1*	No	*RAF1*	Yes	*VWA2*	No
DDB2	Yes	*LEG8*	No	*RAI3*	No	*WAS*	Yes
DDIT3	Yes	*LEPROTL1*	No	*RALGDS*	No	*WBP1*	No
DDIT4	No	*LEU1*	No	*RANB9*	No	*WDR11*	No
DDR1	No	*LFG1*	No	*RANBP2*	Yes	*WDR48*	No
DDR2	Yes	*LGI1*	No	*RAP1A*	No	*WFDC2*	No
DDX10	Yes	*LGR6*	No	*RAP1GDS1*	Yes	*WHSC1*	No
DDX17	No	*LHFP*	No	*RAP2A*	No	*WHSC1L1*	No
DDX3X	Yes	*LHX2*	No	*RAP2B*	No	*WIF1*	Yes
DDX41	No	*LHX4*	No	*RARA*	No	*WIPI1*	No
DDX42	No	*LIFR*	Yes	*RARB*	No	*WIPI2*	No
DDX5	Yes	*LIG1*	No	*RASA1*	No	*WIPI3*	No
DDX6	Yes	*LIG3*	No	*RASF1*	No	*WIPI4*	No
DEC1	No	*LIG4*	No	*RASF2*	No	*WISP1*	No
DEDD	No	*LIMA1*	No	*RASF3*	No	*WN10B*	No
DEDD2	No	*LIMD1*	No	*RASF4*	No	*WNK1*	No
DEF1	No	*LIMK1*	No	*RASF5*	No	*WNK2*	No
DEK	No	*LIN7A*	No	*RASF6*	No	*WNK3*	No
DEMA	No	*LIN9*	No	*RASFA*	No	*WNK4*	No
DENR	No	*LITAF*	No	*RASH*	No	*WNT1*	No
DEP1A	No	*LMBL1*	No	*RASK*	No	*WNT2B*	No
DFFB	No	*LMF1*	No	*RASM*	No	*WNT3*	No

DGAT2	No	*LMNA*	Yes	*RASN*	No	*WNT4*	No
DGCR14	No	*LMO1*	No	*RB*	No	*WNT5A*	No
DGCR8	No	*LMO2*	No	*RB1*	Yes	*WNT6*	No
DHB13	No	*LOX12*	No	*RB11A*	No	*WRN*	No
DHSD	No	*LPP*	No	*RB6I2*	No	*WT1*	No
DI3L2	No	*LRIG3*	No	*RBBP8*	No	*WWOX*	No
DICER	No	*LRP12*	No	*RBCC1*	No	*WWP1*	No
DICER1	Yes	*LRP1B*	Yes	*RBG1L*	No	*WWTR1*	No
DIRA3	No	*LRRN2*	No	*RBL1*	No	*XAB2*	No
DIS3L2	No	*LSM14A*	No	*RBL2*	No	*XAF1*	No
DJC27	No	*LSM7*	No	*RBM10*	Yes	*XBP1*	No
DKC1	No	*LTK*	No	*RBM14*	No	*XIAP*	No
DKK1	No	*LTMD1*	No	*RBM15*	Yes	*XPA*	No
DKK2	No	*LUR1L*	No	*RBM5*	No	*XPC*	No
DKK3	No	*LUZP4*	No	*RBM6*	No	*XPF*	No
DKK4	No	*LYL1*	No	*RBMS1*	No	*XPO1*	Yes
DLEC1	No	*LYN*	No	*RBMX*	No	*XRCC1*	No
DLG1	No	*LYOX*	No	*RBP56*	No	*XRCC2*	No
DLG3	No	*LZTR1*	Yes	*RBTN1*	No	*XRCC3*	No
DLGP5	No	*LZTS1*	No	*RBTN2*	No	*XRCC4*	No
DLP1	No	*LZTS2*	No	*RBX1*	No	*XRCC5*	No
DMAP1	No	*M3K1*	No	*RBX2*	No	*XRCC6*	No
DMBT1	No	*M3K10*	No	*RDM1*	No	*XRN1*	No
DMC1	No	*M3K11*	No	*REC6*	No	*YAF2*	No
DMTF1	No	*M3K14*	No	*RECK*	No	*YAP1*	No
DNAJB1	No	*M3K2*	No	*RECQ4*	No	*YBOX1*	No
DNM2	Yes	*M3K5*	No	*RECQL*	No	*YES*	No
DNMT3A	No	*M3K7*	No	*RECQL4*	Yes	*YETS4*	No
DNTT	No	*M3K8*	No	*RECQL5*	No	*YWHAE*	No
DOCK4	No	*M4K5*	No	*RED*	No	*YYAP1*	No
DOCK8	No	*MAD2L2*	No	*REL*	No	*ZBT16*	No
DP13A	No	*MADD*	No	*REPS2*	No	*ZBT17*	No
DPH1	No	*MAEA*	No	*RET*	No	*ZBT7A*	No
DPM3	No	*MAF*	No	*RETN*	No	*ZBT7C*	No
DPOD1	No	*MAFA*	No	*REV1L*	No	*ZBTB16*	Yes
DPOE1	No	*MAFB*	No	*REV3L*	No	*ZBTB4*	No
DPOLQ	No	*MAFF*	No	*RFC1*	No	*ZC12D*	No
DROSHA	Yes	*MAFG*	No	*RFC2*	No	*ZCCHC8*	No
DUS10	No	*MAFK*	No	*RFC3*	No	*ZDH13*	No
DUS16	No	*MAGI1*	No	*RFC4*	No	*ZDH17*	No
DUS26	No	*MAGI3*	No	*RFC5*	No	*ZEB1*	No
DUS7	No	*MAK*	No	*RFIP3*	No	*ZFHX3*	Yes
DUT	No	*MALAT1*	No	*RFWD3*	No	*ZGPAT*	No
DUX4L1	No	*MALT1*	No	*RGCC*	No	*ZKSC3*	No
E2F1	No	*MAML2*	No	*RGDSR*	No	*ZMAT3*	No
E2F2	No	*MANF*	No	*RGPD3*	No	*ZMY10*	No
E2F3	No	*MAP12*	No	*RGRF1*	No	*ZMY11*	No
E2F6	No	*MAP2K1*	No	*RGS7*	No	*ZMYM3*	No

E2F7	No	*MAP2K2*	No	*RHBDF2*	No	*ZN185*	No
E2F8	No	*MAP2K4*	No	*RHDF2*	No	*ZN217*	No
E41L3	No	*MAP2K7*	No	*RHG07*	No	*ZN320*	No
E4F1	No	*MAP3K1*	No	*RHG20*	No	*ZN350*	No
EBF1	Yes	*MAP3K13*	No	*RHG21*	No	*ZN365*	No
ECT2	No	*MAPK1*	Yes	*RHG26*	No	*ZN513*	No
ECT2L	No	*MAPK3*	No	*RHG29*	No	*ZN521*	No
EED	No	*MAPK5*	No	*RHG35*	No	*ZN655*	No
EFNA1	No	*MAPK8*	No	*RHNO1*	No	*ZN703*	No
EFNA3	No	*MAPK8IP1*	No	*RHOA*	Yes	*ZNF198*	No
EFNA4	No	*MAPK8IP2*	No	*RHOB*	No	*ZNF278*	No
EFNB3	No	*MARE1*	No	*RHOG*	No	*ZNF331*	No
EFS	No	*MARE3*	No	*RHOH*	Yes	*ZNF384*	No
EGF	No	*MAS*	No	*RHXF2*	No	*ZNF429*	No
EGFR	Yes	*MATK*	No	*RIF1*	No	*ZNF479*	No
EGLN1	No	*MATR3*	No	*RIG*	No	*ZNF521*	No
EGLN2	No	*MAVS*	No	*RINT1*	No	*ZNRF3*	No
EGR3	No	*MAX*	No	*RIOX1*	No	*ZRSR2*	No
EHBP1	No	*MB21D2*	No	*RIOX2*	No	*ZSC32*	No
EHF	No	*MBD4*	No	*RIPK1*	No	*ZW10*	No
EI24	No	*MBP*	No	*RIT2*	No	*ZWINT*	No
EIF1AX	No	*MC1R*	No	*RL10*	No	*ZZEF1*	No
EIF3E	Yes	*MCF2*	No	*RL1D1*	No		

Table S2 - Clinicopathological characteristics

Charateristic	Patients, No.(%)	
	With breast cancer	With ovarian cancer
Age at diagnosis, y		
Mean (SD)	41.2 (8.0)	41.0 (14.4)
≤30	4 (9.8)	3 (27.3)
31-45	22 (53.7)	3 (27.3)
≥46	15 (36.6)	5 (45.5)
Breast cancer histology		
Ductal carcinoma in situ	3 (7.3)	NA
Invasive ductal carcinoma	35 (85.4)	NA
Invasive lobular carcinoma	3 (7.3)	NA
Bilateral breast cancer		
Yes	4 (9.7)	NA
No	37 (90.3)	NA
Molecular subtype		
Luminal	27 (65.9)	NA
Triple-negative	7 (17.1)	NA
HER2	3 (7.3)	NA
Without information	4 (9.8)	NA
Ovarian cancer histology		
Serous adenocarcinoma	NA	10 (90.9)
Mucinous adenocarcinoma	NA	1 (9.1)
Second primary cancer		
Without second primary cancer	35 (85.4)	9 (81.8)
Breast	4 (9.8)	2 (18.2)
Melanoma	1 (2.4)	0 (0.0)
Thyroid	1 (2.4)	0 (0.0)

Subtitles:

HER2: human epidermal growth factor receptor-2
NA: not applicable
SD: standard deviation

Table S3 - Personal/family history of cancer

ID sample	Cancer Proband	Diagnosis age	1st degree relative (parents, full siblings, children)	2nd dregree relative (grandparents, grandchildren, aunts, uncles, nephews, nieces or half-siblings)	3rd degree relative (first-cousins, great-grandparents or great grandchildren)
8	Bilateral Breast	51; 52	Breast (F,<69)	Breast (F,36; F,<60; F,<76), Brain (F,?), Leukemia (F,?)	Breast (F,32), Kidney (M,?)
19	Breast	27	Breast (F,<45; F,<45)	Not reported	Not reported
29	Ovarian; Breast	42; 53	Gastric (M,42), Ovarian (F,<60), Uterine (F,57), Breast (F,?)	Uterine (F,?; F,?), Gastric (M,?; M,?; M,?), Ear (M,?), Not specificated (F,?)	Not reported
65	Breast	36	Breast (F,31; F,47)	Breast (F,?; F,?; F,?)	Breast (F,34)
80	Breast	43	Breast (F,?; F,44), Skin (F,?)	Breast (F,60; F,55; F,57), Prostate (M,57), Mediastinum (M,10)	Breast (F,44)
85	Breast	51	Breast (F,45; F,43; F,48), Leukemia (M,?)	Gastric (F,45; M,56)	Not reported
133	Breast	46	Not reported	Uterine (F,?), Lung (F,?), Breast (49), Bone (M,75), Gastric (M,68), Uterine or Ovarian (F,35)	Skin (F,?), Kidney (F,?), Prostate (M,?; M,?), Breast (F,?; F,52; F,56; F,?; F,?)
179	Breast	47	Breast (F,37; F,49; F,61)	Throat (M,?)	Not reported
241	Breast	45	Breast (F,49)	Breast (F,50)	Not reported
256	Breast	47	Not reported	Ovarian (F,50)	Breast (F,36), Colorectal (M,36)
275	Ovarian	60	Prostate (M,80)	Breast (F,32; F,35), Not specificated (M,?; M,?)	Not reported

289	Breast	48	Breast (65)	Breast (F,50)	Breast (F,65)
306	Melanoma; Breast	26; 36	Breast (F,43), Bladder (M,50)	Not reported	Not reported
320	Ovarian	53	Ovarian (F,71), Thyroid (F,29)	Lung (F,83)	Uterine (F,60)
344	Ovarian	47	Prostate (M,64)	Ovarian or Intestine (F,40)	Not reported
361	Breast	56	Bilateral Breast (F,46), Breast (F,42; F,37), Bladder (M,54)	Bladder (M,?)	Bladder (F,?)
426	Breast	38	Breast (F,?)	Colorectal (F,31; F,50)	Breast (F,41)
494	Breast	33	Not reported	Breast (F,38), Ovarian (F,38)	Breast (F,40)
558	Breast	37	Not reported	Breast (F,52), Skin (F,?)	Breast (F,42; F,42; F,46; F,32), Ovarian (F,42)
563	Breast	39	Breast (F,40), Pancreas (F,44)	Not reported	Not reported
565	Ovarian	43	Ovarian (F,?)	Not reported	Not reported
581	Breast	46	Breast (F,46; F,54), Head and Neck (M,83)	Prostate (M,60; M,70)	Not reported
593	Breast	38	Not reported	Not reported	Breast (F,?; F,?; F,?; F,?), Gastric (M,?; M,?), Ovarian (F,?), Colorectal (M,?)
626	Breast	46	Breast (F,74)	Breast (F,57; F,80), Skin (M,?), Multiple myelome (M,60)	Ovarian (F,45), Thyroid (F,40), Pancreas (M,56)
633	Bilateral Breast	38	Gastric (F,29), Duodeum (M,57), Thyroid (F,49), Prostate (M,56)	Colorectal (M,57), Gastric (M,50), Lung (F,50), Not specificated (M,?; M,?; M,?; F,?)	Breast (F,30; F,35; F,40), Gastric (M,?), Prostate (M,?)
638	Breast	42	Breast (F,49; F,50), Gastric (F,55)	Breast (F,?)	Thyroid (F,36), Mouth (F,55)

649	Breast	38	Breast (F,64)	Thyroid (F,64), Ovarian (F,61)	Not reported
656	Ovarian; Breast	41; 48	Lung (M,66)	Ovarian (F,90)	Not reported
689	Bilateral Breast	47	Leukemia (F,47)	Colorectal (M,55), Leukemia (M,?)	Breast (F,48), Colorectal (M,60)
695	Ovarian	21	Not reported	Colorectal (M,40), Breast (F,42)	Not reported
960	Bilateral Breast	59; 70	Uterine (F,45),Breast (F,34), Lung (M,70)	Lung (M,?)	Breast (F,59)
974	Breast	46	Breast (F,55)	Breast (F,60; F,60; F,60; F,65), Prostate (M,70; M,80)	Breast (F,45; F,55), Not specificated (M,50)
1014	Breast	42	Lymphoma (M,19)	Breast (F,53), Melanoma (M,75)	Not reported
1024	Breast	48	Not reported	Gastric (F,41), Breast (F,44; F,?), Ovarian (F,56), Esophagus (F,?), Melanoma (M,?)Not specificated (F,?; F,?)	Prostate (M,56), Colorectal (M,?)
1046	Breast	37	Breast (F,47; F,49)	Not reported	Leukemia (F,40), Colorectal (F,57), Mouth (F,29)
1055	Ovarian	57	Breast (F,50), Pancreas (F,?),	Breast (F,49), Prostate (M,?), Lung (M,?)	Not reported
1095	Breast	43	Breast (F,27; F,42)	Throat (M,72), Uterine (F,98)	Not reported
1096	Breast	46	Skin (F,90), Not specificated (F,64)	Breast (F,44), Prostate (M,50), Lung (M,67), Melanoma (F,50), Colorectal (M,?; M,49; F,75)	Breast (F,?; F,45), Thyroid (F,38)
1097	Bilateral Ovarian	46	Not reported	Not specificated (F,?; F,?)	Breast (F,50)

1109	Breast	21	Not reported	Ovarian (F,43)	Not specificated (M,27)
1151	Breast	38	Breast (F,60)	Breast (F,35)	Not reported
1176	Breast	44	Ovarian (F,30)	Ovarian (F,34), Colon (F,51)	Not reported
1186	Ovarian	20	Not reported	Uterine (F,33)	Leukemia (F,77)
1231	Ovarian	21	Breast (F,43)	Not reported	Not reported
1264	Breast	27	Not reported	Breast (F,50)	Not reported
1294	Breast	26	Breast (F,54)	Breast (F,46; F,35)	Not reported
1299	Breast	50	Breast (F,50; F,50)	Breast (F,?)	Not reported
1308	Breast	45	Breast (F,45)	Pancreas (F,54), Not specificated (F,83)	Breast (F,47)
1326	Breast	35	Not reported	Not reported	Breast (F,26; F,50; F,55)
1327	Breast; Thyroid	37; 39	Breast (F,38)	Breast (F,62)	Leukemia (M,30)
1455	Breast	31	Not reported	Not reported	Breast (F,50; F,52)
1482	Breast	39	Breast (F,42)	Colorectal (M,78),Prostate (M,70), Ovarian (F,54)	Not reported

Subtitle: F: female; M: male

Reference

1. Lu, K.H., Wood, M.E., Daniels, M., Burke, C., Ford, J., Kauff, N.D., Kohlmann, W., Lindor, N.M., Mulvey, T.M., Robinson, L., Rubinstein, W.S., Stoffel, E.M., Snyder, C., Syngal, S., Merrill, J.K., Wollins, D.S., and Hughes, K.S. (2014) American Society of Clinical Oncology Expert Statement: Collection and Use of a Cancer Family History for Oncology Providers. Journal of Clinical Oncology, 32 (8), 833–840.

2. Couch, F.J., Nathanson, K.L., and Offit, K. (2014) Two decades after BRCA: setting paradigms in personalized cancer care and prevention. Science, 343, 1466–70.

3. Fernandes, G.C., Michelli, R.A., Galvao, H.C., Paula, A.E., Pereira, R., Andrade, C.E., Felicio, P.S., Souza, C.P., Mendes, D.R., Volc, S., Berardinelli, G.N., Grasel, R.S., Sabato, C.S., Viana, D.V., Mauad, E.C., Scapulatempo-Neto, C., Arun, B., Reis, R.M., and Palmero, E.I. (2016) Prevalence of BRCA1/BRCA2 mutations in a Brazilian population sample at-risk for hereditary breast cancer and characterization of its genetic ancestry. Oncotarget, 7, 80465–80481.

4. Lu, C., Xie, M., Wendl, M.C., Wang, J., McLellan, M.D., Leiserson, M.D.M., Huang, K., Wyczalkowski, M.A., Jayasinghe, R., Banerjee, T., Ning, J., Tripathi, P., Zhang, Q., Niu, B., Ye, K., Schmidt, H.K., Fulton, R.S., McMichael, J.F., Batra, P., Kandoth, C., Bharadwaj, M., Koboldt, D.C., Miller, C.A., Kanchi, K.L., Eldred, J.M., Larson, D.E., Welch, J.S., You, M., Ozenberger, B.A., Govindan, R., Walter, M.J., Ellis, M.J., Mardis, E.R., Graubert, T.A., Dipersio, J.F., Ley, T.J., Wilson, R.K., Goodfellow, P.J., Raphael, B.J., Chen, F., Johnson, K.J., Parvin, J.D., and Ding, L. (2015) Patterns and functional implications of rare germline variants across 12 cancer types. Nature Communications, 6 (1).

5. Jones, M.R., Kamara, D., Karlan, B.Y., Pharoah, P.D.P., and Gayther, S.A. (2017) Genetic epidemiology of ovarian cancer and prospects for polygenic risk prediction. Gynecologic Oncology, 147 (3), 705–713.

6. Berberich, A.J., Ho, R., and Hegele, R.A. (2018) Whole genome sequencing in the clinic: empowerment or too much information? Canadian Medical Association Journal, 190 (5), E124–E125.

7. Schwarze, K., Buchanan, J., Taylor, J.C., and Wordsworth, S. (2018) Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the literature. GENETICS in MEDICINE.

8. Snape, K., Ruark, E., Tarpey, P., Renwick, A., Turnbull, C., Seal, S., Murray, A., Hanks, S., Douglas, J., Stratton, M.R., and Rahman, N. (2012) Predisposition gene identification in common cancers by exome sequencing: insights from familial breast cancer. Breast Cancer Research and Treatment, 134, 429–433.

9. Cooper DN, Krawczak M, Antonorakis SE (1995) The nature and mechanisms of human gene mutation., in The Metabolic and Molecular Bases of Inherited Disease., Scriver C, Beaudet al., Sly WS, Valle D, editors., New York, pp. 259–291.

10. Gracia-Aznarez, F.J., Fernandez, V., Pita, G., Peterlongo, P., Dominguez, O., de la Hoya, M., Duran, M., Osorio, A., Moreno, L., Gonzalez-Neira, A., Rosa-Rosa, J.M., Sinilnikova, O., Mazoyer, S., Hopper, J., Lazaro, C., Southey, M., Odefrey, F., Manoukian, S., Catucci, I., Caldes, T., Lynch, H.T., Hilbers, F.S.M., van Asperen, C.J., Vasen, H.F.A., Goldgar, D., Radice, P., Devilee, P., and Benitez, J. (2013) Whole Exome Sequencing Suggests Much of Non-BRCA1/BRCA2 Familial Breast Cancer Is Due to Moderate and Low Penetrance Susceptibility Alleles. Plos One, 8.

11. Palmero, E.I., Galvão, H.C.R., Fernandes, G.C., Paula, A.E. de, Oliveira, J.C., Souza, C.P., Andrade, C.E., Romagnolo, L.G.C., Volc, S., C. Neto, M., Sabato, C., Grasel, R., Mauad, E., Reis, R.M., and Michelli, R.A.D. (2016) Oncogenetics service and the Brazilian public health system: the experience of a reference Cancer Hospital. Genetics and Molecular Biology, 39 (2), 168–177.

12. Bolger, A.M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 30 (15), 2114–2120.

13. Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w 1118 ; iso-2; iso-3. Fly, 6 (2), 80–92.

14. Paila, U., Chapman, B.A., Kirchner, R., and Quinlan, A.R. (2013) GEMINI: Integrative Exploration of Genetic Variation and Genome Annotations. PLoS Computational Biology, 9 (7), e1003153.

15. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., Tukiainen, T., Birnbaum, D.P., Kosmicki, J.A., Duncan, L.E., Estrada, K., Zhao, F., Zou, J., Pierce-Hoffman, E., Berghout, J., Cooper, D.N.,

Deflaux, N., DePristo, M., Do, R., Flannick, J., Fromer, M., Gauthier, L., Goldstein, J., Gupta, N., Howrigan, D., Kiezun, A., Kurki, M.I., Moonshine, A.L., Natarajan, P., Orozco, L., Peloso, G.M., Poplin, R., Rivas, M.A., Ruano-Rubio, V., Rose, S.A., Ruderfer, D.M., Shakir, K., Stenson, P.D., Stevens, C., Thomas, B.P., Tiao, G., Tusie-Luna, M.T., Weisburd, B., Won, H.-H., Yu, D., Altshuler, D.M., Ardissino, D., Boehnke, M., Danesh, J., Donnelly, S., Elosua, R., Florez, J.C., Gabriel, S.B., Getz, G., Glatt, S.J., Hultman, C.M., Kathiresan, S., Laakso, M., McCarroll, S., McCarthy, M.I., McGovern, D., McPherson, R., Neale, B.M., Palotie, A., Purcell, S.M., Saleheen, D., Scharf, J.M., Sklar, P., Sullivan, P.F., Tuomilehto, J., Tsuang, M.T., Watkins, H.C., Wilson, J.G., Daly, M.J., MacArthur, D.G., and Exome Aggregation Consortium (2016) Analysis of protein-coding genetic variation in 60,706 humans. Nature, 536, 285.

16. Naslavsky, M.S., Yamamoto, G.L., de Almeida, T.F., Ezquina, S.A.M., Sunaga, D.Y., Pho, N., Bozoklian, D., Sandberg, T.O.M., Brito, L.A., Lazar, M., Bernardo, D.V., Amaro, E., Duarte, Y.A.O., Lebrão, M.L., Passos-Bueno, M.R., and Zatz, M. (2017) Exomic variants of an elderly cohort of Brazilians in the ABraOM database: NASLAVSKY ET AL. Human Mutation, 38 (7), 751–763.

17. Thorvaldsdottir, H., Robinson, J.T., and Mesirov, J.P. (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings in Bioinformatics, 14 (2), 178–192.

18. Forbes, S.A., Tang, G., Bindal, N., Bamford, S., Dawson, E., Cole, C., Kok, C.Y., Jia, M., Ewing, R., Menzies, A., Teague, J.W., Stratton, M.R., and Futreal, P.A. (2010) COSMIC (the Catalogue of Somatic Mutations in Cancer): a resource to investigate acquired mutations in human cancer. Nucleic Acids Res., 38 (Database issue), D652-657.

19. Pletscher-Frankild, S., Pallejà, A., Tsafou, K., Binder, J.X., and Jensen, L.J. (2015) DISEASES: Text mining and data integration of disease–gene associations. Methods, 74, 83–89.

20. The UniProt Consortium (2017) UniProt: the universal protein knowledgebase. Nucleic Acids Res., 45 (D1), D158–D169.

21. Bean, L.J.H., and Hegde, M.R. (2017) Clinical implications and considerations for evaluation of in silico algorithms for use with ACMG/AMP clinical variant interpretation guidelines. Genome Medicine, 9 (1).

22. Kanchi, K.L., Johnson, K.J., Lu, C., McLellan, M.D., Leiserson, M.D.M., Wendl, M.C., Zhang, Q., Koboldt, D.C., Xie, M., Kandoth, C., McMichael, J.F., Wyczalkowski, M.A., Larson, D.E., Schmidt, H.K., Miller, C.A., Fulton, R.S., Spellman, P.T., Mardis, E.R., Druley, T.E., Graubert, T.A., Goodfellow, P.J., Raphael, B.J., Wilson, R.K., and Ding, L. (2014) Integrated analysis of germline and somatic variants in ovarian cancer. Nat Commun, 5, 3156.

23. Ross, J.S., Badve, S., Wang, K., Sheehan, C.E., Boguniewicz, A.B., Otto, G.A., Yelensky, R., Lipson, D., Ali, S., Morosini, D., Chliemlecki, J., Elvin, J.A., Miller, V.A., and Stephens, P.J. (2015) Genomic profiling of advanced-stage, metaplastic breast carcinoma by next-generation sequencing reveals frequent, targetable genomic abnormalities and potential new treatment options. Arch. Pathol. Lab. Med., 139 (5), 642–649.

24. Qian, F., Feng, Y., Zheng, Y., Ogundiran, T.O., Ojengbede, O., Zheng, W., Blot, W., Ambrosone, C.B., John, E.M., Bernstein, L., Hu, J.J., Ziegler, R.G., Nyante, S., Bandera, E.V., Ingles, S.A., Press, M.F., Nathanson, K.L., Hennis, A., Nemesure, B., Ambs, S., Kolonel, L.N., Olopade, O.I., Haiman, C.A., and Huo, D. (2016) Genetic variants in microRNA and microRNA biogenesis pathway genes and breast cancer risk among women of African ancestry. Hum. Genet., 135 (10), 1145–1159.

25. Kawai, S., and Amano, A. (2012) BRCA1 regulates microRNA biogenesis via the DROSHA microprocessor complex. J. Cell Biol., 197 (2), 201–208.

26. Torrezan, G.T., Ferreira, E.N., Nakahata, A.M., Barros, B.D.F., Castro, M.T.M., Correa, B.R., Krepischi, A.C.V., Olivieri, E.H.R., Cunha, I.W., Tabori, U., Grundy, P.E., Costa, C.M.L., de Camargo, B., Galante, P.A.F., and Carraro, D.M. (2014) Recurrent somatic mutation in DROSHA induces microRNA profile changes in Wilms tumour. Nat Commun, 5, 4039.

27. Lynch, H.T., Snyder, C.L., Shaw, T.G., Heinen, C.D., and Hitchins, M.P. (2015) Milestones of Lynch syndrome: 1895–2015. Nature Reviews Cancer, 15, 181.

28. Buckley, A.R., Ideker, T., Carter, H., Harismendy, O., and Schork, N.J. (2018) Exome-wide analysis of bi-allelic alterations identifies a Lynch phenotype in The Cancer Genome Atlas. Genome Med, 10 (1), 69.

29. Couch, F.J., Shimelis, H., Hu, C., and et al (2017) Associations between cancer predisposition testing panel genes and breast cancer. JAMA oncology.

30. Wang, X., Szabo, C., Qian, C., Amadio, P.G., Thibodeau, S.N., Cerhan, J.R., Petersen, G.M., Liu, W., and Couch, F.J. (2008) Mutational analysis of thirty-two double-strand DNA break repair genes in breast and pancreatic cancers. Cancer Res., 68 (4), 971–975.

31. Brandalize, A.P.C., Schüler-Faccini, L., Hoffmann, J.-S., Caleffi, M., Cazaux, C., and Ashton-Prolla, P. (2014) A DNA repair variant in POLQ (c.-1060A > G) is associated to hereditary breast cancer patients: a case-control study. BMC Cancer, 14, 850.

32. Family, L., Bensen, J.T., Troester, M.A., Wu, M.C., Anders, C.K., and Olshan, A.F. (2015) Single-nucleotide polymorphisms in DNA bypass polymerase genes and association with breast cancer and breast cancer subtypes among African Americans and Whites. Breast Cancer Res. Treat., 149 (1), 181–190.

33. Lu, H.-M., Li, S., Black, M.H., Lee, S., Hoiness, R., Wu, S., Mu, W., Huether, R., Chen, J., Sridhar, S., Tian, Y., McFarland, R., Dolinsky, J., Tippin Davis, B., Mexal, S., Dunlop, C., and Elliott, A. (2018) Association of Breast and Ovarian Cancers With Predisposition Genes Identified by Large-Scale Sequencing. JAMA Oncology.

34. Kurian, A.W., Hughes, E., Handorf, E.A., Gutin, A., Allen, B., Hartman, A.-R., and Hall, M.J. (2017) Breast and Ovarian Cancer Penetrance Estimates Derived From Germline Multiple-Gene Sequencing Results in Women. JCO Precision Oncology, (1), 1–12.

35. Connor, A.A. (2017) Germline & somatic variation in gastrointestinal malignancies.

36. Penkert, J., Schmidt, G., Hofmann, W., Schubert, S., Schieck, M., Auber, B., Ripperger, T., Hackmann, K., Sturm, M., Prokisch, H., Hille-Betz, U., Mark, D., Illig, T., Schlegelberger, B., and Steinemann, D. (2018) Breast cancer patients suggestive of Li-Fraumeni syndrome: mutational spectrum, candidate genes, and unexplained heredity. Breast Cancer Research, 20 (1).

37. El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., Qureshi, M., Richardson, L.J., Salazar, G.A., Smart, A., Sonnhammer, E.L.L., Hirsh, L., Paladin, L., Piovesan, D., Tosatto, S.C.E., and Finn, R.D. (2019) The Pfam protein families database in 2019. Nucleic Acids Research, 47 (D1), D427–D432.

38. Song, H., Dicks, E., Ramus, S.J., Tyrer, J.P., Intermaggio, M.P., Hayward, J., Edlund, C.K., Conti, D., Harrington, P., Fraser, L., Philpott, S., Anderson, C., Rosenthal, A., Gentry-Maharaj, A., Bowtell, D.D., Alsop, K., Cicek, M.S., Cunningham, J.M., Fridley, B.L., Alsop, J., Jimenez-Linan, M., Høgdall, E., Høgdall, C.K., Jensen, A., Kjaer, S.K., Lubiński, J., Huzarski, T.,

Jakubowska, A., Gronwald, J., Poblete, S., Lele, S., Sucheston-Campbell, L., Moysich, K.B., Odunsi, K., Goode, E.L., Menon, U., Jacobs, I.J., Gayther, S.A., and Pharoah, P.D.P. (2015) Contribution of Germline Mutations in the RAD51B, RAD51C, and RAD51D Genes to Ovarian Cancer in the Population. J. Clin. Oncol., 33 (26), 2901–2907.

39. Fan, C., Zhang, J., Ouyang, T., Li, J., Wang, T., Fan, Z., Fan, T., Lin, B., and Xie, Y. (2018) RAD50 germline mutations are associated with poor survival in BRCA1/2-negative breast cancer patients. Int. J. Cancer, 143 (8), 1935–1942.

6.2. Candidate DNA repair susceptibility genes identified by whole-exome sequencing in non-*BRCA1*/*BRCA2* carriers at high risk for hereditary breast/ovarian cancer

Paula Silva Felicio[1]; Rebeca Silveira Grasel[1]; Natalia Campacci[1]; Henrique de Campos Reis Galvão[2]; Giovana Tardin Torrezan[3]; Andre Escremim de Paula[4]; Cristina da Silva Sabato[4]; Gabriela Carvalho Fernandes[4]; Cristiano de Pádua Souza[2]; Rodrigo Depieri Michelli[2]; Carlos Eduardo Andrade[2]; Marcus de Medeiros Matsushita[5]; Timothée Revil[6]; Jiannis Ragoussis[6,7]; Fergus J. Couch[8,9]; Steven N. Hart[8,9]; Rui M. Reis[1,4,10,11]; Patricia N. Tonin[5,12,13]; Matias Eliseo Melendez[1]; Dirce Maria Carraro[3,14]; Edenir Inêz Palmero[1,4,15]

[1]Molecular Oncology Research Center, Barretos Cancer Hospital, Barretos, Brazil.

[2]Department of Oncogenetics, Barretos Cancer Hospital, Barretos, Brazil.

[3]Genomics and Molecular Biology Group, CIPE - A. C. Camargo Cancer Center, São Paulo, Brazil.

[4]Center of Molecular Diagnosis, Barretos Cancer Hospital, Barretos, São Paulo, Brazil.

[5] Department of Pathology, Barretos Cancer Hospital, Barretos, Brazil

[6]Department of Human Genetics, McGill University, Montreal, Canada.

[7]McGill Genome Centre, University of McGill, Montreal, Canada.

[8]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota.

[9]Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota.

[10]Life and Health Sciences Research Institute (ICVS), Medical School, University of Minho, Braga, Portugal.

[11]ICVS/3B's-PT Government Associate Laboratory, Braga/Guimarães, Portugal.

[12]Department of Medicine, McGill University, Montreal, Canada.

[13]Cancer Research Program, The Research Institute of the McGill University Health Centre, Montreal, Canada.

[14]Genomic Diagnostic Center, AC Camargo Cancer Center, São Paulo, Brazil.

[15]Barretos School of Health Sciences, Dr. Paulo Prata - FACISB, Barretos, Brazil.

Keywords: BRCAX, DNA repair genes, whole-exome sequencing, hereditary breast and ovarian cancer predisposition syndrome, familial cancer.

Abstract

Germline pathogenic *BRCA1* or *BRCA2* variants are responsible for approximately 25% of hereditary breast/ovarian cancer (HBOC) cases. Other than *BRCA1/BRCA2*, several DNA repair genes have been associated with HBOC. However, for approximately 50% of families at-risk for HBOC, the causal factor is not known. Thus, the present study aimed to search for germline variants in DNA repair genes of different pathways. We evaluated, through whole-exome sequencing (WES) of germline DNA, variants involved in DNA repair pathways in 52 Brazilian families at high risk for HBOC that previously tested negative for *BRCA1/BRCA2/TP53* pathogenic variants. To select the most promising variants, we used function-based variant prioritization, information from population databases (gnomAD and AbraOM) and *in silico* prediction tools (CADD scaled, REVEL and M-CAP). Then, all variants considered to be damaging by the bioinformatics tools were classified according to the ACMG-AMP guidelines and ClinVar. Variants classified as benign/probably benign were excluded. We identified a total of 29 (55.8%) patients with germline variants in 26 different DNA repair genes. Using the ACMG-AMP guidelines, four variants classified as likely-pathogenic/pathogenic were identified (*CHEK2*, *FAN1*, *PMS2* and *RAD51C* genes). These variants were identified in families with breast cancer (*CHEK2*, *FAN1* and *PMS2*) and ovarian cancer (*RAD51C*). Another important gene emphasized in our cases was *RAD54L*, which was observed in two women with a personal and family history of ovarian cancer. Our results highlight the fundamental role of the DNA repair pathways in hereditary cancer development and the importance of including those genes in genetic testing for at-risk patients.

Introduction

The pathways associated with DNA repair are a primary protection system characterized by a phenomenal multienzyme and are responsible for ensuring the integrity of the cellular genome[153]. Every human being is regularly exposed to damaging metabolic products, environmental chemicals and radiation that harm DNA, which can consequently alter the genetic information. Furthermore, spontaneous errors during DNA replication can cause DNA damage[154]. However, all genes involved in DNA repair pathways can provide genome integrity through repair of any detected damage[153].

Germline variants in DNA repair genes are the underlying cause of many hereditary cancer predisposition syndromes, such as hereditary breast and ovarian cancer (HBOC)[128,155,156], Lynch syndrome[157,158(p2)], Fanconi anemia (FA)[159], and Xeroderma Pigmentosum[160]. Regarding HBOC syndrome, the most frequent genes altered are *BRCA1* and *BRCA2*, which contribute to approximately 20-25% of cases[128,161]. Pathogenic variants of these genes lead to defective proteins and compromise the function of the homologous repair (HR) pathway. In addition, other genes involved in DNA repair pathways are associated with breast/ovarian cancer development in non-*BRCA1/BRCA2* patients, such as *ATM*[162], *CHEK2*[163], *PALB2*[63], *RAD51C*[164] and *TP53*[165]. However, for a large proportion of HBOC patients (50%), the genetic cause associated with the development of breast and ovarian cancers is unknown[19,86,165]. Thus, these patients do not have ideal clinical management, and preventive and risk-reducing measures become limited.

In light of these issues, we conducted this study to investigate the presence of germline variants in DNA repair pathway genes by whole-exome sequencing (WES) in 52 Brazilian non-*BRCA1/BRCA2* women diagnosed with breast and/or ovarian cancer.

Materials and Methods

Ethical criteria

This study was approved by the ethics committee of the Barretos Cancer Hospital (BCH) (40814115.4.0000.5437). All patients signed a written informed consent form.

Study samples and whole-exome sequencing

The patient cohort consisted of 52 unrelated Brazilian women at high risk for HBOC who were diagnosed with breast/ovarian cancer. These patients were referred for genetic testing from the Oncogenetics Department of BCH, as described elsewhere[166]. All patients previously tested negative for *BRCA1/BRCA2* and *TP53* germline pathogenic variants.

Details about the methodology, variant calling process and filters are described elsewhere (Felicio *et al.* manuscript under evaluation), but briefly, for selecting the most promising variants, we used function-based variant prioritization. For quality filtering, variants with vertical coverage ≥10x and variant allele frequency (VAF) ≥0.25 were selected. Next, a total of 228 DNA repair genes, published by Das and Ghosh (2017)[167], were analyzed. Variants

present in the population-based Genome Aggregation Database (gnomAD[168]) and the Brazilian database of WES from 609 healthy individuals (AbraOM—Brazilian genomic variants[169]) that had a frequency ≤2% (minor allele frequency [MAF] ≤0.02) were maintained. In addition, to evaluate the potential impact of the pathogenicity of missense variants, we used three bioinformatic tools, CADD (score ≥20), REVEL (score ≥0.5) and M-CAP (score ≥0.025). Furthermore, all variants detected after applied filters were classified according to the ACMG-AMP guidelines[170,171] as pathogenic, likely-pathogenic, with uncertain significance, and likely-benign or benign and were compared to the ClinVar deposits (Figure 1).

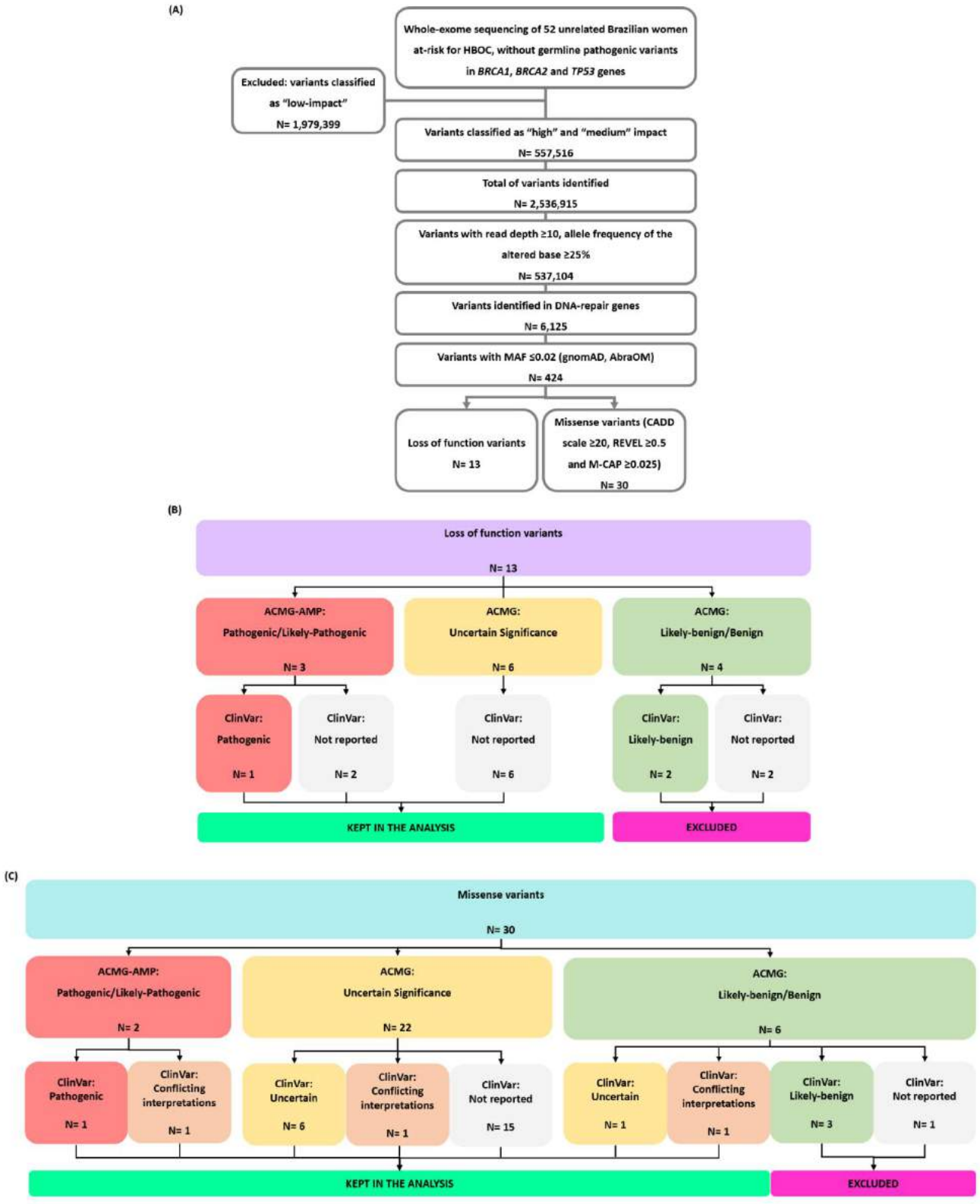


Figure 1 - Schematic representation of variant prioritization in WES-involved DNA repair genes. WES data from 52 unrelated Brazilian women at risk for HBOC, without germline pathogenic variants in *BRCA1*, *BRCA2* and *TP53* genes. **(A)** Variants classified as "high-impact" and "medium-impact" by snpEFF/GEMINI were prioritized. Next, variants with base coverage

≥10x and variant allele frequency (VAF) ≥0.25 were selected. Subsequently, variants present in population databases with frequency ≤2% (MAF ≤0.02) were maintained for analysis. The missense variants with a CADD scale ≥20, a REVEL ≥0.5 and an M-CAP ≥0.025 were maintained. **(B)** Classification of the loss of function variants identified in our cases using ACMG-AMP criteria and the ClinVar database. **(C)** Classification of the missense variants identified in our cases using ACMG-AMP criteria and the ClinVar database.

Results

1. Patient diagnosis

A WES dataset of 52 unrelated women at risk for HBOC without pathogenic *BRCA1/BRCA2* and *TP53* variants was filtered to select uncommon missense and loss of function (LoF) variants implicated in DNA repair pathways. As described elsewhere (Felicio *et al.* manuscript under evaluation), a total of 41 women had breast cancer (BC), nine had ovarian cancer (OC), and two were diagnosed with both OC and BC.

We identified 29 patients (55.8%) with germline variants classified as rare (MAF ≤0.02) and potentially damaging using *in silico* tools (scores: CADD ≥20, REVEL ≥0.5 and M-CAP ≥0.025) involving DNA repair pathways. Of these, 22 had BC, five had OC, and two had a personal history of OC and BC. Moreover, the average age of cancer diagnosis was 40.6 years, ranging from 21 to 56 years.

2. Germline variants in DNA repair genes

The schematic representation of variant prioritization is shown in Figure 1. Next, we implemented variant prioritization and aimed to evaluate the clinical significance. We also compared the variants by ACMG-AMP criteria and ClinVar deposits. Variants described by ACMG-AMP and ClinVar as benign/likely-benign were removed from our analysis (Supplementary Table 1). Thus, in our final data, we maintained a total of 35 variants (nine LoF and 26 missense) that occurred in 29 women (55.8%) and affected 26 different DNA-repair genes (Figure 2).

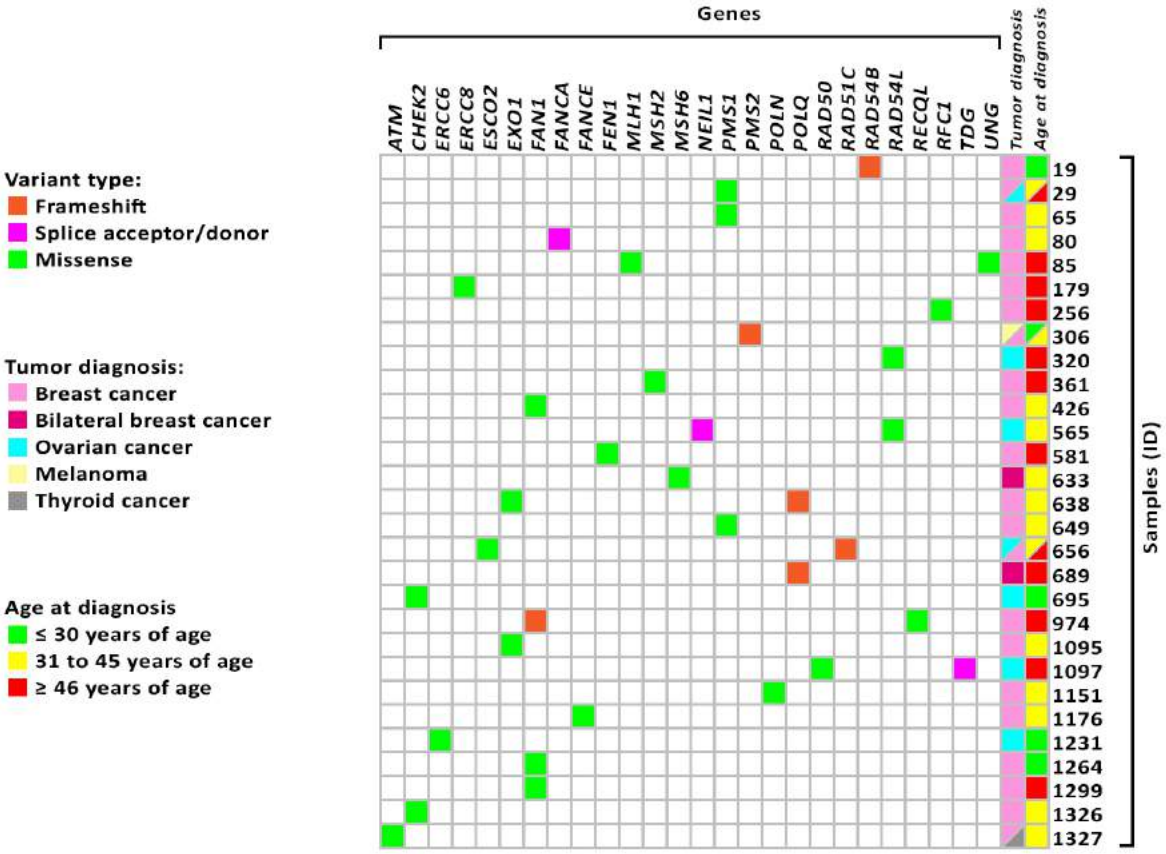


Figure 2 - Rare variants identified involving DNA repair genes. In green: missense; in purple: splice acceptor/donor; in orange: frameshift. Information about tumor diagnosis: in light pink: breast cancer; in dark pink: bilateral breast cancer; in light blue: ovarian cancer; in light yellow: melanoma; in gray: thyroid cancer. Age at diagnosis: in green: diagnosis ≤ 30 years of age; in yellow: 31 to 45 years of age; in red: ≥ 46 years of age is presented.

To better understand the potential underlying mechanisms and the most commonly altered pathways implicated in non-*BRCA1/BRCA2* mutation carriers, we analyzed network interactions involving DNA repair genes from two publicly available databases (KEGG and STRING). As a result of the network-based analysis, we observed a total of 185 interactions involving the 26 genes (Figure 3).

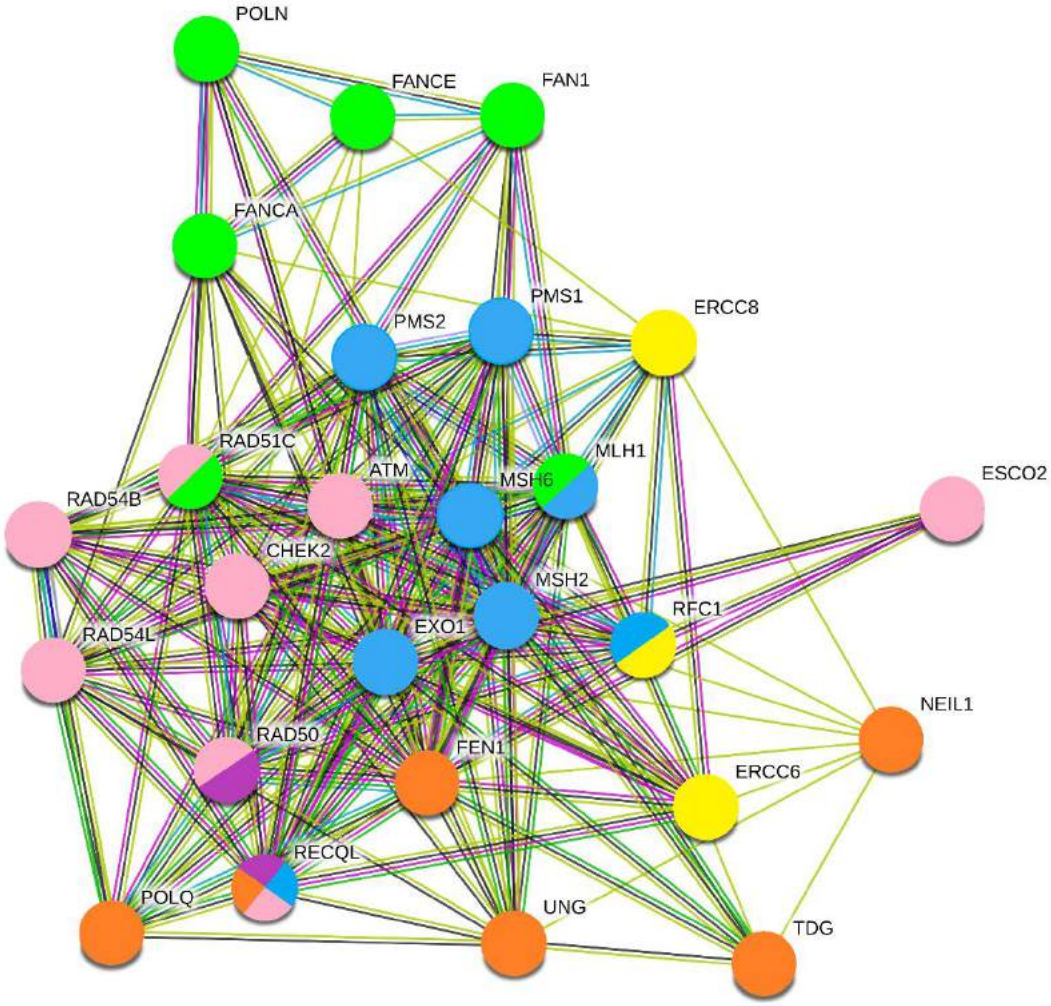


Figure 3 – STRING pathway overview of DNA repair genes identified in 29 women at high risk for hereditary breast and ovarian cancer. Legends: in light pink: genes involved in homologous recombination, in green: genes involved Fanconi anemia, in blue: genes involved in mismatch repair, in orange: genes involved in base excision repair, in purple: genes involved in non-homologous end-joining, and in yellow: genes involved in nucleotide excision repair.

According to KEGG and STRING and as depicted in Figure 3, we found that the most altered pathway was homologous recombination, followed by mismatch repair, Fanconi anemia, base excision repair, nucleotide excision repair and non-homologous end-joining.

Information about the pathways, genes, and variants identified in our study is discussed below. In addition, according to the ACMG-AMP criteria and ClinVar database, we separated the variants into different categories. Variants considered benign/likely-benign by

ClinVar and ACMG-AMP are described in Supplementary Table 1 and were excluded from the analysis and comparisons performed. In addition, 29 variants of unknown clinical significance (by ACMG-AMP and/or ClinVar) were identified and are described in Table 1. Finally, those variants considered pathogenic/likely-pathogenic are listed in Table 2.

Table 1 – Rare variants identified in DNA-repair and classified as uncertain significance by ACMG-AMP.

ID sample	Diagnosis (Age)	BC - Family history	OC - Family history	Gene	Pathway	cDNA variant – HGVS	Protein level	ACMG-AMP	ClinVar	Transcript
19	Breast (27)	Yes	No	*RAD54B*	HR	c.2733delG	p.(Ter911Tyrext*15)	Uncertain significance	Not reported	ENST00000336148
29	Ovarian (42), Breast (53)	Yes	Yes	*PMS1*	MMR	c.856G>A	p.(Gly286Arg)	Uncertain significance	Not reported	ENST00000409593
65	Breast (36)	Yes	No	*PMS1*	MMR	c.1625G>A	p.(Arg542His)	Uncertain significance	Not reported	ENST00000409593
80	Breast (43)	Yes	No	*FANCA*	FA	c.435+2dupT	.	Uncertain significance	Not reported	ENST00000561722
85	Breast (51)	Yes	No	*MLH1*	MMR/FA	c.794G>A	p.(Arg265His)	Uncertain significance	Uncertain significance	ENST00000231790
				UNG	BER	c.262C>T	p.(Arg88Cys)	Uncertain significance	Uncertain significance	ENST00000242576
179	Breast (47)	Yes	No	*ERCC8*	NER	c.839C>A	p.(Thr280Lys)	Uncertain significance	Conflicting interpretations†	ENST00000265038
256	Breast (47)	Yes	Yes	*RFC1*	MMR/NER	c.2014G>A	p.(Val672Met)	Uncertain significance	Not reported	ENST00000349703
320	Ovarian (53)	No	Yes	*RAD54L*	HR	c.604C>T	p.(Arg202Cys)	Uncertain significance	Not reported	ENST00000371975
361	Breast (56)	Yes	No	*MSH2*	MMR	c.80C>T	p.(Pro27Leu)	Uncertain significance	Uncertain significance	ENST00000233146
426	Breast (38)	Yes	No	*FAN1*	FA	c.149T>G	p.(Met50Arg)	Uncertain significance	Not reported	ENST00000362065
565	Ovarian (43)	No	Yes	*NEIL1*	BER	c.434+2T>C	.	Uncertain significance	Not reported	ENST00000355059
				RAD54L	HR	c.1094G>A	p.(Arg365Gln)	Uncertain significance	Not reported	ENST00000371975
581	Breast (46)	Yes	No	*FEN1*	BER	c.311G>A	p.(Arg104Gln)	Uncertain significance	Not reported	ENST00000305885
633	Bilateral Breast (38)	Yes	No	*MSH6*	MMR	c.2885T>C	p.(Ile962Thr)	Uncertain significance	Uncertain significance	ENST00000234420
638	Breast (42)	Yes	No	*POLQ*	BER	c.4262_4268delTACTATT	p.(Ile1421ArgfsTer8)	Uncertain significance	Not reported	ENST00000264233
				EXO1	MMR	c.820G>A	p.(Gly274Arg)	Uncertain significance	Not reported	ENST00000348581
649	Breast (38)	Yes	Yes	*PMS1*	MMR	c.856G>A	p.(Gly286Arg)	Uncertain significance	Not reported	ENST00000409593

656	Ovarian (41), Breast (48)	No	Yes	*ESCO2*	HR	c.1735C>A	p.(Pro579Thr)	Uncertain significance	Uncertain significance	ENST00000305188
689	Bilateral Breast (47)	Yes	No	*POLQ*	BER	c.4262_4268delTACTATT	p.(Ile1421ArgfsTer8)	Uncertain significance	Not reported	ENST00000264233
695	Ovarian (21)	Yes	No	*CHEK2*	HR	c.538C>T	p.(Arg180Cys)	Likely-pathogenic	Conflicting interpretations‡	ENST00000328354
974	Breast (46)	Yes	No	*RECQL*	BER/MMR/NHEJ/HR	c.401C>T	p.(Thr134Ile)	Uncertain significance	Uncertain significance	ENST00000314748
1095	Breast (43)	Yes	No	*EXO1*	MMR	c.820G>A	p.(Gly274Arg)	Uncertain significance	Not reported	ENST00000348581
1097	Bilateral Ovarian (46)	Yes	No	*TDG*	BER	c.1090_1090+1insTTGAGAGC	p.(Val367Leufs*6)	Uncertain significance	Not reported	ENST00000392872
				RAD50	HR/NHEJ	c.353T>C	p.(Ile118Thr)	Likely-benign	Uncertain significance	ENST00000265335
1151	Breast (38)	Yes	No	*POLN*	FA	c.2133T>G	p.(Phe711Leu)	Uncertain significance	Not reported	ENST00000382865
1176	Breast (44)	No	Yes	*FANCE*	FA	c.329C>T	p.(Pro110Leu)	Uncertain significance	Not reported	ENST00000229769
1231	Ovarian (21)	Yes	No	*ERCC6*	NER	c.1801G>A	p.(Gly601Ser)	Uncertain significance	Not reported	ENST00000355832
1264	Breast (27)	Yes	No	*FAN1*	FA	c.149T>G	p.(Met50Arg)	Uncertain significance	Not reported	ENST00000362065
1299	Breast (50)	Yes	No	*FAN1*	FA	c.2524T>C	p.(Tyr842His)	Uncertain significance	Not reported	ENST00000362065
1327	Breast (37), Thyroid (39)	Yes	No	*ATM*	HR	c.1595G>A	p.(Cys532Tyr)	Likely-benign	Conflicting interpretations§	ENST00000278616

Subtitles: BC: breast cancer; BER: Base Excision Repair; FA: Fanconi Anemia; HR: Homologous Recombination; MMR: Mismatch Repair; NER: Nucleotide Excision Repair; NHEJ: Non-Homologous End-Joining; OC: Ovarian cancer; †: likely-benign(2); uncertain significance(2) - acess at 10/05/2019; ‡: benign(3); likely-benign(2); uncertain significance(7) - acess at 10/05/2019; §: likely-benign(2); uncertain significance(6) - acess at 10/05/2019.

Table 2 – Rare variants identified in DNA-repair and classified as likely-pathogenic/pathogenic by ACMG-AMP.

ID sample	Diagnosis (Age)	BC - Family history	OC - Family history	Gene	Pathway	cDNA variant - HGVS	Protein level	ACMG-AMP	ClinVar	Transcript
306	Melanoma (26), Breast (36)	Yes	No	*PMS2*	MMR	c.2182_2184delACTinsG	p.(Thr728AlafsTer7)	Pathogenic	Pathogenic	ENST00000265849
656	Ovarian (41), Breast (48)	No	Yes	*RAD51C*	HR/FA	c.890_899delTTGTTCCTGC	p.(Leu297HisfsTer2)	Pathogenic	Not reported	ENST00000337432
974	Breast (46)	Yes	No	*FAN1*	FA	c.357_358delGG	p.(Glu120SerfsTer10)	Likely-pathogenic	Not reported	ENST00000362065
1326	Breast (35)	Yes	No	*CHEK2*	HR	c.349A>G	p.(Arg117Gly)	Likely-pathogenic	Pathogenic/Likely-pathogenic	ENST00000328354

Subtitles: BC: Breast cancer; FA: Fanconi Anemia; HR: Homologous Recombination; MMR: Mismatch Repair; OC: Ovarian cancer.

2.1 Homologous Recombination (HR)

The most significantly altered pathway was HR. We observed eight different genes altered, including the RAD-family genes (*RAD50*, *RAD51C*, *RAD54B*, *RAD54L*), *ATM*, *CHEK2*, *ESCO2* and *RECQL*. We found two variants classified by ACMG-AMP/ClinVar as likely-pathogenic/pathogenic involving *RAD51C* (c.890_899delTTGTTCCTGC) and *CHEK2* (c.349A>G). Regarding the variant in *RAD51C*, the patient (ID 656) was diagnosed with OC followed by BC (41 and 48 years of age, respectively) and reported the presence of OC in the family history (Supplementary Figure 1). Regarding the variant in *CHEK2*, the patient (ID 1326) was diagnosed with BC at 35 years of age and reported the presence of BC in the family history (Supplementary Figure 2).

Additionally, the majority of the other variants identified in this pathway were classified as Class 3 by ACMG-AMP, which involved *CHEK2* (c.538C>T), *RAD54B* (c.2733delG), *RAD54L* (c.604C>T and c.1094G>A), *ESCO2* (c.1735C>A) and *RECQL* (c.401C>T). However, it is important to draw attention to the *RAD54L* gene. Both variants identified in this gene showed a higher score of pathogenicity using *in silico* tools (CADD and REVEL), although it has not been reported by ClinVar. Both index cases (ID samples: 320 and 565) *RAD54L* mutation carrier (c.604C>T and c.1094G>A, respectively) developed OC and reported the presence of OC in a first-degree relative and no cases of BC were diagnosed in either family (Supplementary Figure 3A and Supplementary Figure 3B). Regarding the variants identified in *ATM* (c.1595G>A) and *RAD50* (c.353T>C), it is important to emphasize that these variants were classified as likely-benign by ACMG-AMP and showed a discrepancy of interpretation by ClinVar, which were classified as likely-benign/uncertain significance and having uncertain significance, respectively.

2.2 Mismatch Repair (MMR)

The other most significantly altered pathway was MMR, with a total of eight different genes altered (*EXO1*, *MLH1*, *MSH2*, *MSH6*, *PMS1*, *PMS2*, *RECQL* and *RFC1*). Regarding the LoF variants, we found a frameshift in the *PMS2* gene (c.2182_2184delACTinsG). This variant was considered pathogenic by ACMG-AMP and ClinVar. This variant in *PMS2* was identified in a patient (ID 306) with melanoma and BC (26 and 36 years of age, respectively) who reported the presence of BC in the family history. Moreover, through segregation analysis, we observed the same variant (c.2182_2184delACTinsG) was present in all three tested members (sisters),

however, only one member was diagnosed with two primary tumors: BC (at 42 years of age) followed by thyroid cancer (at 43 years of age) (Supplementary Figure 4).

Regarding missense variants, most of them were classified as Class 3 (uncertain significance) by ACMG-AMP and affected the following genes: *EXO1* (c.820G>A), *MLH1* (c.794G>A), *MSH2* (c.80C>T), *MSH6* (c.2885T>C), *PMS1* (c.856G>A and c.1625G>A), *RECQL* (c.401C>T) and *RFC1* (c.2014G>A) (Table 1). It is important to emphasize that the variants identified in *MLH1* and *RFC1* showed a higher score of pathogenicity using *in silico* tools (CADD and REVEL).

2.3 Fanconi Anemia (FA)

The other pathway most significantly altered was FA (Figure 3), totaling six different genes affected (*FANCA*, *FANCE*, *FAN1*, *MLH1*, *POLN* and *RAD51C*). In this pathway, we found nine different variants, which included three LoF variants and six missense variants. Of these, the variants identified at the *FAN1* (c.357_358delGG) and *RAD51C* gene (c.890_899delTTGTTCCTGC) were considered likely-pathogenic and pathogenic, respectively, according to ACMG-AMP criteria and not reported by ClinVar. Regarding the LoF variant in *FAN1*, the patient (ID 974) was diagnosed with BC at 46 years of age and reported the presence of BC in the family history (Supplementary Figure 5A). About the personal and family history of cancer involving the patient with *RAD51C*, we discussed above in "*2.1 Homologous Recombination*". Furthermore, six variants (*FANCA*: c.435+2dup; *FANCE*: c.329C>T; *FAN1*: c.149T>G and c.2524T>C; *MLH1*: c.794G>A; and *POLN*: c.2133T>G) were considered to have uncertain significance by ACMG-AMP. Moreover, only the variant observed in *MLH1* (c.794G>A) was reported by ClinVar as having uncertain significance, and the remaining variants were not found in ClinVar. Still, two patients (ID 426 and 1264) shared a missense variant identified in *FAN1*, c.149T>G. These patients had a personal and family history of BC (Supplementary Figure 5B and Supplementary Figure C). In addition, the patient ID 426 reported the presence of two relatives diagnosed with colorectal cancer.

2.4 Base excision repair (BER)

Following the greatest significantly altered pathway, we found germline variants in genes involved in the single-strand DNA repair pathway BER. As detailed in Table 1, a total of six different genes were affected. Of these, three genes showed LoF variants (*NEIL1*:

c.434+2T>C; *POLQ*: c.4262_4268delTACTATT and *TDG*: c.1090_1090+1insTTGAGAGC) and three missense variants (*FEN1*: c.311G>A; *RECQL*: c.401C>T and *UNG*: c.262C>T). Regarding the ACMG-AMP criteria, all variants were classified as having uncertain significance, and only two variants (in *RECQL* and *UNG*) were reported by ClinVar (with the same classification). The remaining variants were not found in ClinVar.

2.5 Nucleotide excision repair (NER)

Although at a low frequency, we identified genes associated with the NER pathway. Three different genes were found in this pathway: *ERCC6* (*c.1801G>A*), *ERCC8* (c.839C>A) and *RFC1* (c.2014G>A). All variants found were missense and classified as Class 3 according to ACMG-AMP criteria. Moreover, only the variant observed in *ERCC8* showed conflicting interpretations of pathogenicity by ClinVar (likely-benign/uncertain significance).

2.6 Non-Homologous End-Joining (NHEJ)

Alterations in genes involved in the NHEJ pathway were also observed in the patients. We identified two altered genes involved in this pathway, both of which had missense variants. Of the variants, we highlight *RAD50* (c.353T>C) and *RECQL* (c.401C>T) – both described above in "*2.1 Homologous Recombination*".

Discussion

The purpose of our study was to identify candidate genes involved in DNA repair pathways to explain the personal/family history of breast/ovarian cancer development. For this, we performed WES in constitutive DNA in a cohort of familial breast/ovarian patients compatible with autosomal dominant inheritance and without pathogenic variants in the *BRCA1/BRCA2* and *TP53* genes.

DNA repair pathways are intimately associated with many cellular processes, such as DNA replication and recombination, cell cycle checkpoint and apoptosis. Moreover, defects that impair the function of DNA repair genes can result in the presence of genetic variants and genomic instability[154]. Several studies have associated specific DNA repair pathways with different cancers, such as MMR in colorectal cancer and HR in breast cancer[172–175]. Here, we report six DNA repair pathways with potentially pathogenic germline variants enriched in our cohort.

Of the 52 women at high risk for HBOC, we identified a total of 29 (55.8%) patients with germline variants in 26 different DNA repair genes. In addition, we observed that four patients had a pathogenic/likely-pathogenic variant (6.9%) and 27 (51.91%) had an uncertain significance variant by ACMG-AMP classification. Regarding the diagnosis of these women, 22 had BC, five had OC, and two had a personal history of OC and BC. In addition, the average age at cancer diagnosis was 40.6 years, ranging from 21 to 56 years. Regarding the variants, a total of 35 variants (nine loss of function and 26 missense) were observed. Moreover, a high frequency of potentially pathogenic germline variants was found in the HR, MMR and FA pathways, which will be discussed below.

Among the HR pathway, two likely-pathogenic/pathogenic variants by ACMG-AMP/ClinVar, including *RAD51C* (also involved in the FA pathway) and *CHEK2*, were observed. We found a LoF variant in *RAD51C* that has been shown to be associated with an increased risk for OC. Some studies show that this gene can represent a moderate to high increase in the risk of OC (OR: 4.98 [95% CI: 3.09-8.04][62] and OR: 5.2 [95% CI: 1.1-24.0][102]). Furthermore, according to NCCN guidelines, the lifetime risk of OC in *RAD51C*-mutation carriers seems to be sufficient to justify consideration of risk-reducing salpingo-oophorectomy (RRSO)[176]. However, studies showed that *RAD51C* was not related to increased risk of BC[177] (OR: 0.78 [95% CI: 0.47-1.37][86] and OR: 1.43 [95% CI: 0.97-2.12][62]). Regarding *CHEK2* gene, we identified a pathogenic variant (c.349A>G) involving this gene, which has already been associated with increased risk of BC (OR: 2.26 [95% CI: 1.89-2.72][86]). A recent study published by Kleiblova *et al.* (2019), through analysis of *CHEK2* in 1,928 patients at high risk for HBOC, identified 10 LoF variants in 46 (2.39%) patients and 26 missense variants in 88 (4.56%) patients. Nevertheless, through *in silico* analysis, the authors inferred that the presence of rare and damaging variants in *CHEK2* had a significantly increased BC risk (OR: 3.90 [95% CI: 1.24-13.35]) and OC risk (OR: 4.77 [95% CI: 0.77-22.47])[178]. Furthermore, a study published by Southey *et al.* (2016) aimed to estimate some variants associated with cancer risks. Of these variants, one was the same variant that we found in our cohort (c.349A>G and c.538C>T), and it was analyzed in this study. The authors found evidence of an association with BC risk for both variants in *CHEK2*. The authors estimated that the variant c.349A>G had an OR of 2.26 (95% CI: 1.29-3.95), and the other variant c.538C>T showed an OR of 1.33 (95% CI: 1.05- 1.67)[179]. Regarding our cases, the woman with the variant c.349A>G was diagnosed with BC at 35 years of age, and the other patient with the variant c.538C>T was diagnosed with OC at 21 years of age. In addition, both

women reported the presence of BC cases in the family history. Additionally, according to NCCN guidelines, an annual mammogram is indicated for clinical management. However, evidence of a connection between *CHEK2* carriers and OC development was insufficient[176]. Still on the HR pathway, we highlight two missense variants involving *RAD54L* (c.604C>T and c.1094G>A), both observed in patients with personal/family history of OC. Although the function of *RAD54L* in HBOC syndrome still not clear, some authors reported the presence of germline variants in diagnosed with BC and OC. Matsuda *et al.* (1999) identified the presence of somatic variant (c.973G>A) in a patient diagnosed with BC at 63 years of age and without a family history of cancer. Moreover, the authors observed the absence of the wildtype allele in the normal tissue, indicated that this variant was a germline variant and this gene has suppressor tumor features. In addition, the authors also reported that this variant was absent in 100 individuals without cancer[180]. Shagimardanova *et al.* (2018)[181], through data published in the Annals of Oncology, analyzed 568 patients diagnosed with BC/OC, between 21 to 82 years of age. In addition, the authors reported that 193/568 (34%) patients had clinical features of HBOC syndrome. Regarding the variants identified, 128 (22.5%) was carried pathogenic or likely-pathogenic variants in the *BRCA1* gene; 52 (9.2%) in *BRCA2* and, 100 (17.6%) in one of other HBOC related genes, including *RAD54L*. More details about the variants in non-*BRCA* genes were not published by the authors, however, they highlight that the use of multigene panel is a powerful approach to find variants in other genes, which could play a role in tumorigenesis.

The other pathway that deserves attention is the MMR, which has been shown to increase susceptibility to BC, OC, or both[62,86,165], although available evidence is conflicting. We highlight the LoF variant in *PMS2* (c.2182_2184delACTinsG), which was observed in a patient diagnosed with melanoma (at 26 years of age) followed by BC (at 36 years of age). Moreover, we observed the variant segregation, which was present in all three tested members (sisters). Important to emphasize that one sister had a diagnosis of two primary tumors: BC (at 42 years of age) followed by thyroid cancer (at 43 years of age). Regarding *PMS2*, most literature reports showed that this gene was not associated with increased BC risk (OR: 0.89 [95% CI: 0.66-1.22][62] and 0.82 [95% CI: 0.44-1.47][86]). Additionally, the NCCN guidelines report insufficient evidence for evaluating BC risk associated with *PMS2* variants[176]. However, Roberts *et al.* (2018), through a case-control study, reported an association between pathogenic variants in *PMS2* and BC in women who were not *BRCA1/BRCA2* mutation carriers.

Due to the high frequency of germline variants in *PMS2*, the authors concluded that this gene can be associated with an increased risk for BC and should be considered for women who have a personal/family history of BC[182]. Additionally, LaDuca *et al.* (2019) observed that 37.4% of patients with a pathogenic variant in *PMS2* had clinical criteria for the indication of genetic testing for *BRCA1/BRCA2*. Thus, the authors highlighted the importance of using the multigene panel in patients at risk for hereditary cancer, even when patients had a classic clinical feature of a particular syndrome, resulting in treatment and follow-up opportunities for families[113].

Regarding the FA pathway, we underline the *FAN1* gene, in which a total of four patients showed a germline variant and all cases had a personal and family history of BC. However, only one variant was LoF and was described as likely-pathogenic by ACMG-AMP (c.357_358delGG). Moreover, two patients (ID 426 and 1264) shared a missense variant identified in *FAN1* (c.149T>G). Both patients had a personal and family history of BC. Additionally, the patient ID 426 reported the presence of two relatives diagnosed with colorectal cancer. The variant c.149T>G was previously reported by Smith *et al.* (2016) in patients with a personal and family history of pancreatic cancer. In addition, the authors observed the co-segregation of the variant into two different families, thus considering a candidate gene for pancreatic cancer predisposition[183]. Based on the above considerations, Lachaud *et al.* (2016) through *in vitro* assays, aimed to evaluate the impact of variant c.149T>G. The authors reported that due to variant localization in a functional domain (UBZ: ubiquitin binding domain), there was a loss of protein function. Therefore, the researchers concluded that the c.149T> G variant was associated with the development of pancreatic cancer[184]. Some authors have been reported the possible association of germline variants in the *FAN1* gene with colorectal cancer[185,186], although the reported variants are not the same ones identified in our study. However, few studies reported the possible association of germline variants in *FAN1* and the high risk for BC development. Park *et al.* (2011)[187], through WES of early-onset BC families from Australia, identified two different variants involving *FAN1* gene (c.1129C>T and c.1520G>A). The authors observed that the variant c.1129C>T was present in all six tested members with cancer (five with BC and one with melanoma). In addition, the authors genotyped both variants in 1.417 population-based cases and 1.490 unaffected population-based controls. The authors concluded that these variants were rare in the Australian population (MAF of 0.0064 and 0.010, respectively) and were not related with BC risk (OR: 0.80, 95% CI [0.39 to 1.61], OR: 0.74, 95% CI [0.41 to 1.29], respectively).

Therefore, the authors concluded that future studies need to be performed to better elucidate the possible role of the *FAN1* gene as a breast cancer susceptibility gene.

Another enriched pathway was BER, in which we observed six different genes altered. We found three LoF variants (*NEIL1*, *POLQ* and *TDG*) and four missense variants (*FEN1*, *RECQL* and *UNG*) classified as having uncertain significance according to AMCG-AMP criteria. Few studies have reported the presence of germline variants in these genes in non-*BRCA1/BRCA2*-mutation carriers. However, Kanchi *et al.* (2014) reported the presence of LoF variants in *NEIL1* and *UNG* in women diagnosed with OC[188]. Moreover, some authors have proposed that deficiency of the BER and HR pathways together leads to synthetically lethal death[189,190] but can benefit from PARP inhibitors. In our cohort, we observed two patients with variants involved in the BER and HR pathways. One of them (ID 565), who had a personal/family history of OC, showed a LoF variant in *NEIL1* and a missense variant in *RAD54L*. Another patient (ID 1097), who had a personal history of OC, showed a LoF variant in *TDG* and a missense variant in *RAD50*. Regarding *POLQ*, few studies have reported the presence of germline variants. A study published by Stafford *et al.* (2017), through WES in 48 patients with OC, identified the presence of a LoF variant[191]. In addition, Ceccaldi *et al.* (2015) reported that olaparib (Lynparza) may also be effective in treating patients diagnosed with breast and ovarian tumors that express abnormally high levels of POLQ. The authors described a synthetic lethal relationship between the HR pathway and *POLQ*-mediated repair in ovarian tumors, suggesting that POLQ may be a potential target for cancer therapy[192].

In addition, we found germline variants in other pathways, such as NHEJ and NER. Regarding NHEJ, we found two different genes altered, *RAD50* and *RECQL*. Both variants were reported as having uncertain significance by ClinVar. However, the variant identified in *RAD50* (c.353T>C) was classified as likely benign by ACMG-AMP, and the variant observed in *RECQL* (c.401C>T) was classified as having uncertain significance by ACMG-AMP. We did not find germline variants in the literature involving the majority of these genes in patients at risk of HBOC. Regarding the NER pathway, we identified two genes altered, *ERCC6* and *ERCC8*. Germline variants involved in the *ERCC* genes have been associated with xeroderma pigmentosum, which increases the risk for several cancers[193]. Kanchi *et al.* (2014), through tumor analysis of OC, identified a somatic truncation variant in *ERCC8*[188]. However, we did not find, in the literature, the presence of germline variants in these genes or a possible association with BC/OC in cases diagnosed at an early age.

The current study represents one of the largest investigations of women at high risk of HBOC without pathogenic variants in *BRCA1*/*BRCA2*/*TP53* that aimed to identify rare germline variants associated with DNA repair pathways. Some limitations of our study include the absence of analysis of noncoding regions and large genomic rearrangements.

Conclusion

In summary, we detailed a characterization of germline variants in putative DNA repair genes using WES in Brazilian non-*BRCA1/BRCA2* and *TP53*-carrier women diagnosed with BC/OC. Our findings suggest that some DNA repair genes, such as *CHEK2*, *FAN1*, *PMS2*, *RAD51C* and *RAD54L*, may play a role in non-*BRCA1/BRCA2* patients and may be predisposing factors in a portion of HBOC cases. Additionally, the majority of the variants (82.9%) identified in our study were described as having uncertain significance by ACMG-AMP and not reported by ClinVar (65.7%). Large-scale studies such as this can help with the molecular identification of germline variants in non-*BRCA1/BRCA2/TP53* patients and provide extra evidence for the association of moderate-risk genes with breast/ovarian cancer development, contributing to more effective screening and targeted therapies.

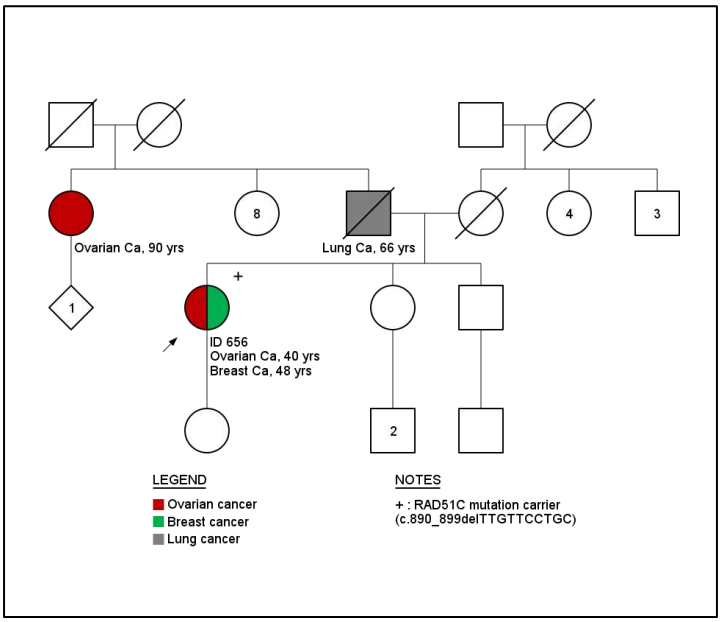
Acknowledgments

We are very thankful to all patients who participated in this study. This project is funded through grants from FAPESP (2013/24633-2), PRONON (25000.056766/2015-64), Coordination for the Improvement of Higher Education Personnel (CAPES/PROSUP) and McGill University (Department of Medicine and Cancer Research Program). EIP and RMR are recipients of CNPq Productivity Grants.

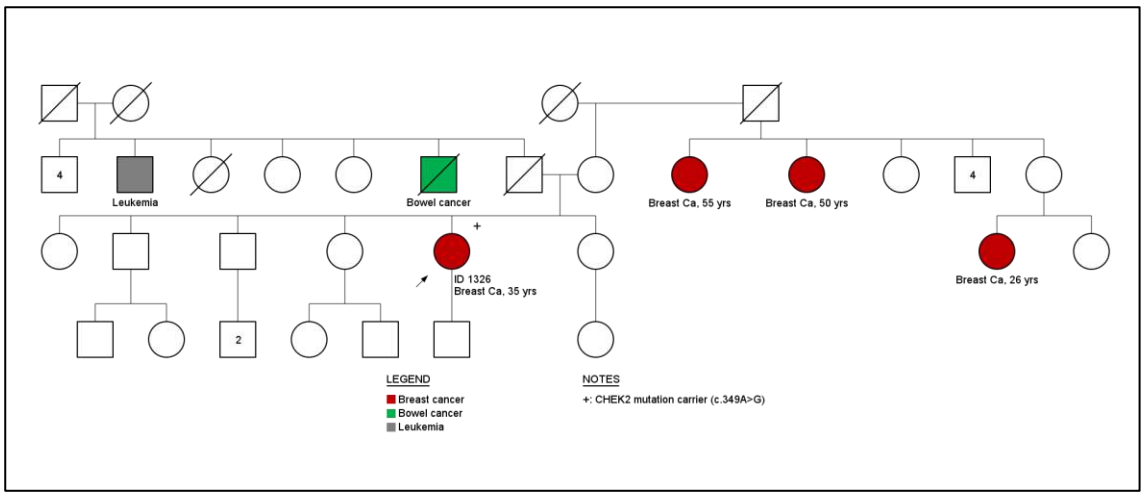
Conflict of Interest

The authors have nothing to disclose.

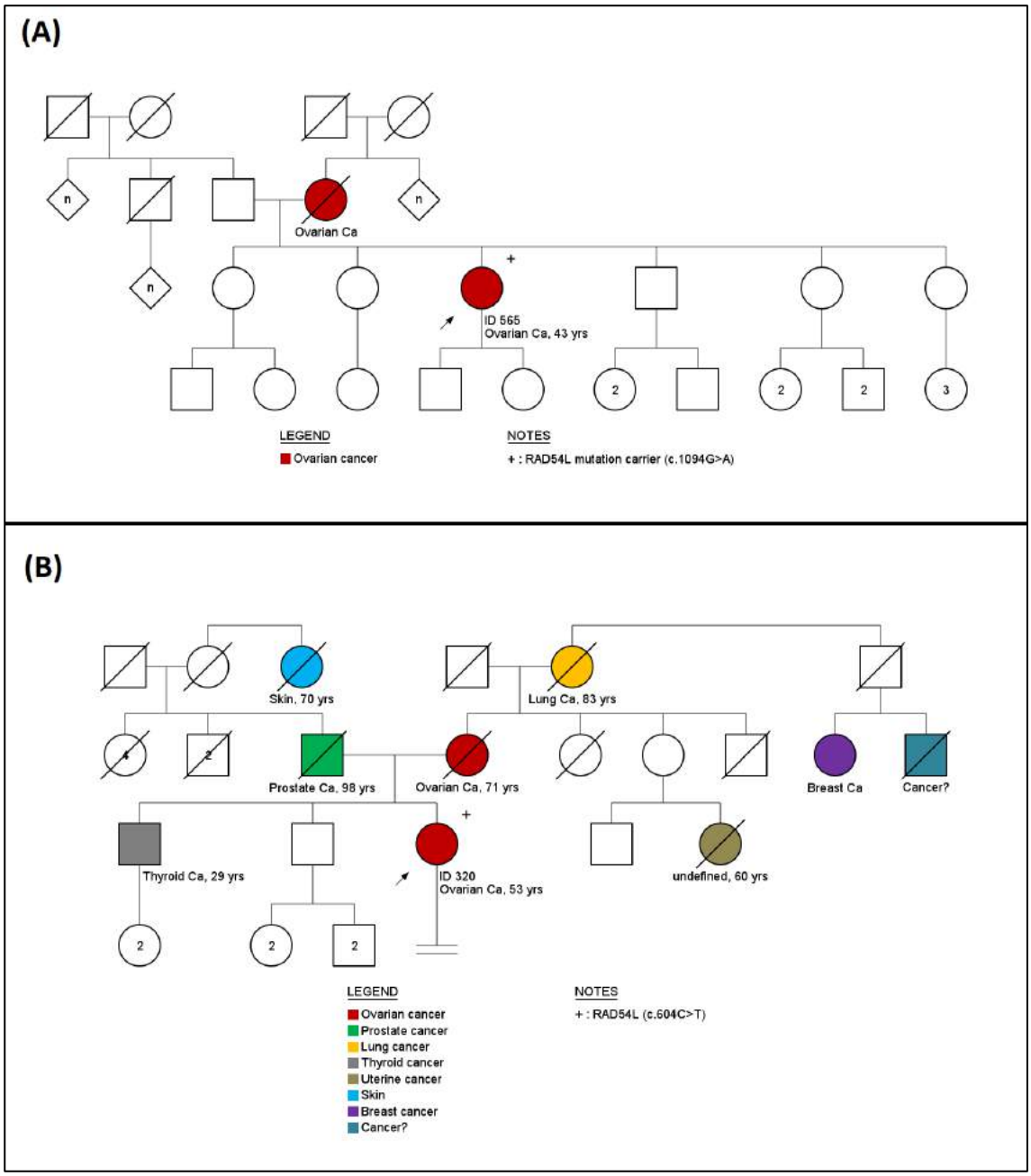
SUPPLEMENTARY FIGURES



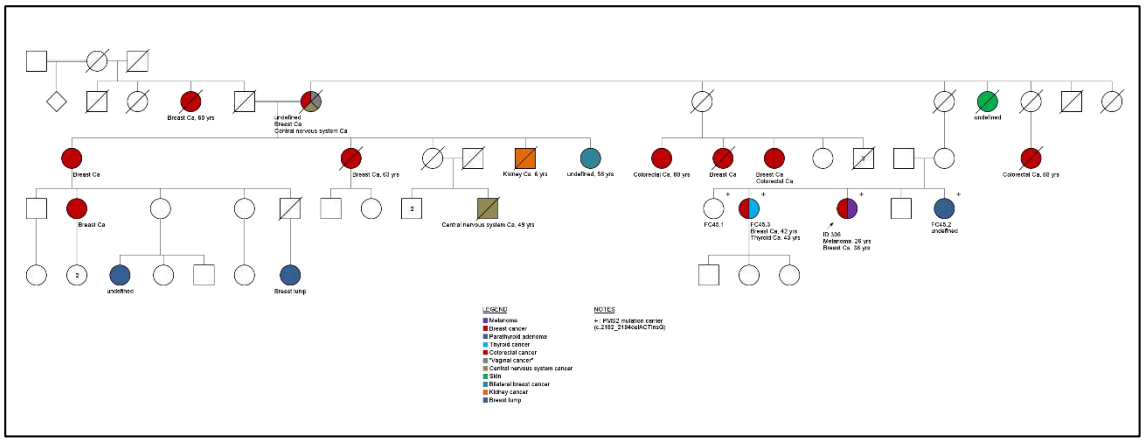
Supplementary Figure 1 - Pedigree of index case (ID 656), *RAD51C* mutation carrier.



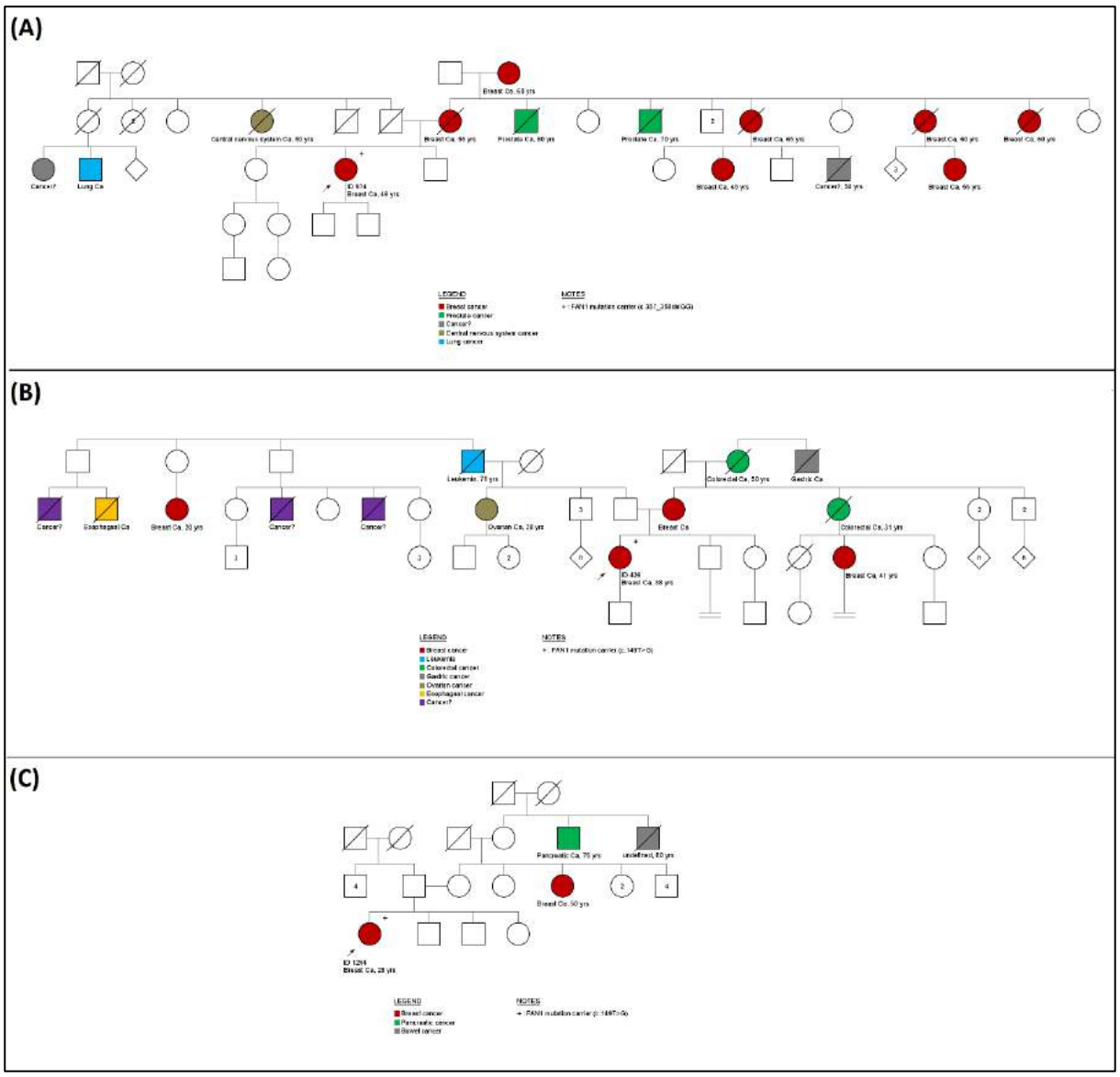
Supplementary Figure 2 - Pedigree of index case (ID 1326), *CHEK2* mutation carrier.



Supplementary Figure 3 - Pedigree of index case **(A)**: ID 565, *RAD54L* (c.1094G>A) mutation carrier. **(B)**: ID 320, *RAD54L* (c.604C>T) mutation carrier.



Supplementary Figure 4 - Pedigree of index case (ID 306), *PMS2* mutation carrier.



Supplementary Figure 5 - Pedigree of index case *FAN1* mutation carriers (A): ID 974, (B): ID 426, (C): ID 1264.

SUPPLEMENTARY TABLE 1

Supplementary Table 1 – Rare variants identified in DNA-repair and classified as likely-benign/benign by ACMG-AMP (excluded variants).

ID sample	Diagnosis (Age)	BC - Family history	OC - Family history	Gene	Pathway	cDNA variant - HGVS	Protein level	ACMG-AMP	ClinVar	Transcript
85	Breast (51)	Yes	No	*RECQL4*	HR	c.1847A>G	p.(Asn616Ser)	likely-benign	likely-benign	ENST00000428558
133	Breast (46)	Yes	Yes	*MLH1*	MMR/FA	c.1852_1853delAAinsGC	p.(Lys618Ala)	benign	benign	ENST00000231790
320	Ovarian (53)	No	Yes	*NHEJ1*	NHEJ	c.589-17151A>T	.	benign	not reported	ENST00000483627
				MSH2	MMR	c.23C>T	p.(Thr8Met)	likely-benign	benign	ENST00000233146
593	Breast (38)	Yes	Yes	*ERCC1*	NER/FA	c.875G>A	p.(Trp292Ter)	likely-benign	likely-benign	ENST00000013807
656	Ovarian (41), Breast (48)	No	Yes	*ATM*	HR	c.5558A>T	p.(Asp1853Val)	likely-benign	likely-benign	ENST00000278616
1055	Ovarian (57)	Yes	No	*MSH5*	MMR	c.1129C>T	p.(Leu377Phe)	likely-benign	not reported	ENST00000375703
1186	Ovarian (20)	No	No	*POLM*	NHEJ	c.1431delG	p.(Lys478SerfsTer25)	benign	not reported	ENST00000395831

References

1. Curtin NJ. DNA repair dysregulation from cancer driver to therapeutic target. Nature Reviews Cancer. 2012; 12: 801–17. doi: 10.1038/nrc3399.

2. Torgovnick A, Schumacher B. DNA repair mechanisms in cancer development and therapy. Frontiers in Genetics [Internet]. 2015 [cited 2019 Feb 24]; 6. doi: 10.3389/fgene.2015.00157.

3. Aloraifi F. Detection of novel germline mutations for breast cancer in non-BRCA1/2 families. FEBS Journal. 2016; 283: 5. doi: 10.1111/febs.13803.

4. Torrezan GT, de Almeida FG dos SR, Figueiredo MCP, Barros BD de F, de Paula CAA, Valieris R, de Souza JES, Ramalho RF, da Silva FCC, Ferreira EN, de Nóbrega AF, Felicio PS, Achatz MI, et al. Complex Landscape of Germline Variants in Brazilian Patients With Hereditary and Early Onset Breast Cancer. Frontiers in Genetics [Internet]. 2018 [cited 2019 Mar 12]; 9. doi: 10.3389/fgene.2018.00161.

5. Nielsen FC, van Overeem Hansen T, Sorensen CS. Hereditary breast and ovarian cancer: new genes in confined pathways. Nature reviews Cancer. 2016; 16: 599–612. doi: 10.1038/nrc.2016.72.

6. Esteban-Jurado C, Vila-Casadesús M, Garre P, Lozano JJ, Pristoupilova A, Beltran S, Muñoz J, Ocaña T, Balaguer F, López-Cerón M, Cuatrecasas M, Franch-Expósito S, Piqué JM, et al. Whole-exome sequencing identifies rare pathogenic variants in new predisposition genes for familial colorectal cancer. Genet Med. 2015; 17: 131–42. doi: 10.1038/gim.2014.89.

7. Ten Broeke SW, van Bavel TC, Jansen AML, Gómez-García E, Hes FJ, van Hest LP, Letteboer TGW, Olderode-Berends MJW, Ruano D, Spruijt L, Suerink M, Tops CM, van Eijk R, et al. Molecular Background of Colorectal Tumors From Patients With Lynch Syndrome Associated With Germline Variants in PMS2. Gastroenterology. 2018; 155: 844–51. doi: 10.1053/j.gastro.2018.05.020.

8. Schneider M, Chandler K, Tischkowitz M, Meyer S. Fanconi anaemia: genetics, molecular biology, and cancer - implications for clinical management in children and adults. Clinical Genetics. 2015; 88: 13–24. doi: 10.1111/cge.12517.

9. Weon JL, Glass DA. Novel therapeutic approaches to xeroderma pigmentosum. British Journal of Dermatology [Internet]. 2018 [cited 2019 Feb 25]; . doi: 10.1111/bjd.17253.

10. Neamatzadeh H, Shiryazdi SM, Kalantar SM. BRCA1 and BRCA2 mutations in Iranian breast cancer patients: A systematic review. J Res Med Sci. 2015; 20: 284–93.

11. Slavin TP, Maxwell KN, Lilyquist J, Vijai J, Neuhausen SL, Hart SN, Ravichandran V, Thomas T, Maria A, Villano D, Schrader KA, Moore R, Hu C, et al. Author Correction: The contribution of pathogenic variants in breast cancer susceptibility genes to familial breast cancer risk. npj Breast Cancer. 2017; 3: 44. doi: 10.1038/s41523-017-0046-2.

12. Mavaddat N, Michailidou K, Dennis J, Lush M, Fachal L, Lee A, Tyrer JP, Chen T-H, Wang Q, Bolla MK, Yang X, Adank MA, Ahearn T, et al. Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. Am J Hum Genet. 2019; 104: 21–34. doi: 10.1016/j.ajhg.2018.11.002.

13. Antoniou AC, Casadei S, Heikkinen T, Barrowdale D, Pylkas K, Roberts J, Lee A, Subramanian D, De Leeneer K, Fostira F, Tomiak E, Neuhausen SL, Teo ZL, et al. Breast-cancer risk in families with mutations in PALB2. The New England journal of medicine. 2014; 371: 497–506. doi: 10.1056/NEJMoa1400382.

14. Janatova M, Soukupova J, Stribrna J, Kleiblova P, Vocka M, Boudova P, Kleibl Z, Pohlreich P. Mutation Analysis of the RAD51C and RAD51D Genes in High-Risk Ovarian Cancer Patients and Families from the Czech Republic. Brusgaard K, editor. PLoS ONE. 2015; 10: e0127711. doi: 10.1371/journal.pone.0127711.

15. Lu H-M, Li S, Black MH, Lee S, Hoiness R, Wu S, Mu W, Huether R, Chen J, Sridhar S, Tian Y, McFarland R, Dolinsky J, et al. Association of Breast and Ovarian Cancers With Predisposition Genes Identified by Large-Scale Sequencing. JAMA Oncology [Internet]. 2018 [cited 2018 Nov 10]; . doi: 10.1001/jamaoncol.2018.2956.

16. Couch FJ, Shimelis H, Hu C, et al. Associations between cancer predisposition testing panel genes and breast cancer. JAMA oncology. 2017; . doi: 10.1001/jamaoncol.2017.0424.

17. Jones MR, Kamara D, Karlan BY, Pharoah PDP, Gayther SA. Genetic epidemiology of ovarian cancer and prospects for polygenic risk prediction. Gynecologic Oncology. 2017; 147: 705–13. doi: 10.1016/j.ygyno.2017.10.001.

18. Palmero EI, Galvão HCR, Fernandes GC, Paula AE de, Oliveira JC, Souza CP, Andrade CE, Romagnolo LGC, Volc S, C. Neto M, Sabato C, Grasel R, Mauad E, et al. Oncogenetics service and the Brazilian public health system: the experience of a reference Cancer Hospital. Genetics and Molecular Biology. 2016; 39: 168–77. doi: 10.1590/1678-4685-gmb-2014-0364.

19. Das R, Ghosh SK. Genetic variants of the DNA repair genes from Exome Aggregation Consortium (EXAC) database: significance in cancer. DNA Repair. 2017; 52: 92–102. doi: 10.1016/j.dnarep.2017.02.013.

20. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, Tukiainen T, Birnbaum DP, Kosmicki JA, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016; 536: 285.

21. Naslavsky MS, Yamamoto GL, de Almeida TF, Ezquina SAM, Sunaga DY, Pho N, Bozoklian D, Sandberg TOM, Brito LA, Lazar M, Bernardo DV, Amaro E, Duarte YAO, et al. Exomic variants of an elderly cohort of Brazilians in the ABraOM database: NASLAVSKY ET AL. Human Mutation. 2017; 38: 751–63. doi: 10.1002/humu.23220.

22. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL, ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015; 17: 405–24. doi: 10.1038/gim.2015.30.

23. Kalia SS, Adelman K, Bale SJ, Chung WK, Eng C, Evans JP, Herman GE, Hufnagel SB, Klein TE, Korf BR, McKelvey KD, Ormond KE, Richards CS, et al. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. Genet Med. 2017; 19: 249–55. doi: 10.1038/gim.2016.190.

24. Heeke AL, Pishvaian MJ, Lynce F, Xiu J, Brody JR, Chen W-J, Baker TM, Marshall JL, Isaacs C. Prevalence of Homologous Recombination-Related Gene Mutations Across Multiple Cancer Types. JCO Precis Oncol. 2018; 2018. doi: 10.1200/PO.17.00286.

25. Romero-Laorden N, Castro E. Inherited mutations in DNA repair genes and cancer risk. Current Problems in Cancer. 2017; 41: 251–64. doi: 10.1016/j.currproblcancer.2017.02.009.

26. AlDubayan SH, Giannakis M, Moore ND, Han GC, Reardon B, Hamada T, Mu XJ, Nishihara R, Qian Z, Liu L, Yurgelun MB, Syngal S, Garraway LA, et al. Inherited DNA-Repair Defects in Colorectal Cancer. The American Journal of Human Genetics. 2018; 102: 401–14. doi: 10.1016/j.ajhg.2018.01.018.

27. den Brok WD, Schrader KA, Sun S, Tinker AV, Zhao EY, Aparicio S, Gelmon KA. Homologous Recombination Deficiency in Breast Cancer: A Clinical Review. JCO Precision Oncology. 2017; : 1–13. doi: 10.1200/PO.16.00031.

28. Kurian AW, Hughes E, Handorf EA, Gutin A, Allen B, Hartman A-R, Hall MJ. Breast and Ovarian Cancer Penetrance Estimates Derived From Germline Multiple-Gene Sequencing Results in Women. JCO Precision Oncology. 2017; : 1–12. doi: 10.1200/PO.16.00066.

29. Song H, Dicks E, Ramus SJ, Tyrer JP, Intermaggio MP, Hayward J, Edlund CK, Conti D, Harrington P, Fraser L, Philpott S, Anderson C, Rosenthal A, et al. Contribution of Germline Mutations in the RAD51B , RAD51C , and RAD51D Genes to Ovarian Cancer in the Population. Journal of Clinical Oncology. 2015; 33: 2901–7. doi: 10.1200/JCO.2015.61.2408.

30. Daly MB, Pilarski R, Berry M, Buys SS, Farmer M, Friedman S, Garber JE, Kauff ND, Khan S, Klein C, Kohlmann W, Kurian A, Litton JK, et al. NCCN Guidelines Insights: Genetic/Familial High-Risk Assessment: Breast and Ovarian, Version 2.2017. Journal of the National Comprehensive Cancer Network : JNCCN. 2017; 15: 9–20.

31. Lee K, Seifert BA, Shimelis H, Ghosh R, Crowley SB, Carter NJ, Doonanco K, Foreman AK, Ritter DI, Jimenez S, Trapp M, Offit K, Plon SE, et al. Clinical validity assessment of genes frequently tested on hereditary breast and ovarian cancer susceptibility sequencing panels. Genet Med. 2019; 21: 1497–506. doi: 10.1038/s41436-018-0361-5.

32. Kleiblova P, Stolarova L, Krizova K, Lhota F, Hojny J, Zemankova P, Havranek O, Vocka M, Cerna M, Lhotova K, Borecka M, Janatova M, Soukupova J, et al. Identification of deleterious germline CHEK2 mutations and their association with breast and ovarian cancer. Int J Cancer. 2019; 145: 1782–97. doi: 10.1002/ijc.32385.

33. Southey MC, Goldgar DE, Winqvist R, Pylkäs K, Couch F, Tischkowitz M, Foulkes WD, Dennis J, Michailidou K, van Rensburg EJ, Heikkinen T, Nevanlinna H, Hopper JL, et al. PALB2 , CHEK2 and ATM rare variants and cancer risk: data from COGS. J Med Genet. 2016; 53: 800–11. doi: 10.1136/jmedgenet-2016-103839.

34. Matsuda M, Miyagawa K, Takahashi M, Fukuda T, Kataoka T, Asahara T, Inui H, Watatani M, Yasutomi M, Kamada N, Dohi K, Kamiya K. Mutations in the RAD54 recombination gene in primary cancers. Oncogene. 1999; 18: 3427–30. doi: 10.1038/sj.onc.1202692.

35. Shagimardanova E, Shigapova LH, Enikeev RF, Druzhkov M, Nikitin A, Gusev O, Gordiev M. 240PLandscape of germline mutations in hereditary breast and ovarian cancer (HBOC) patients in Russia revealed by target panel sequencing. Annals of Oncology [Internet]. 2018 [cited 2019 Oct 4]; 29. doi: 10.1093/annonc/mdy270.235.

36. Roberts ME, Jackson SA, Susswein LR, Zeinomar N, Ma X, Marshall ML, Stettner AR, Milewski B, Xu Z, Solomon BD, Terry MB, Hruska KS, Klein RT, et al. MSH6 and PMS2 germ-line pathogenic variants implicated in Lynch syndrome are associated with breast cancer. Genetics in Medicine. 2018; 20: 1167–74. doi: 10.1038/gim.2017.254.

37. LaDuca H, Polley EC, Yussuf A, Hoang L, Gutierrez S, Hart SN, Yadav S, Hu C, Na J, Goldgar DE, Fulk K, Smith LP, Horton C, et al. A clinical guide to hereditary cancer panel testing: evaluation of gene-specific cancer associations and sensitivity of genetic testing criteria in a cohort of 165,000 high-risk patients. Genet Med. 2019; . doi: 10.1038/s41436-019-0633-8.

38. Smith AL, Alirezaie N, Connor A, Chan-Seng-Yue M, Grant R, Selander I, Bascuñana C, Borgida A, Hall A, Whelan T, Holter S, McPherson T, Cleary S, et al. Candidate DNA repair susceptibility genes identified by exome sequencing in high-risk pancreatic cancer. Cancer Lett. 2016; 370: 302–12. doi: 10.1016/j.canlet.2015.10.030.

39. Lachaud C, Moreno A, Marchesi F, Toth R, Blow JJ, Rouse J. Ubiquitinated Fancd2 recruits Fan1 to stalled replication forks to prevent genome instability. Science. 2016; 351: 846–9. doi: 10.1126/science.aad5634.

40. Vargas G, Dámaso E, Navarro M, Pons T, Mina L, Fernández A, Feliubadaló L, Solanes A, Iglesias S, Velasco À, Balmaña J, Ramón y Cajal T, Valencia A, et al. Identification of germline FAN1 variants in MSH2-deficient Lynch-like syndrome patients. 2016; . Available from

https://ehtg.org/2016/04/p19-identification-germline-fan1-variants-msh2-deficient-lynch-like-syndrome-patients/

41. Seguí N, Mina LB, Lázaro C, Sanz-Pamplona R, Pons T, Navarro M, Bellido F, López-Doriga A, Valdés-Mas R, Pineda M, Guinó E, Vidal A, Soto JL, et al. Germline Mutations in FAN1 Cause Hereditary Colorectal Cancer by Impairing DNA Repair. Gastroenterology. 2015; 149: 563–6. doi: 10.1053/j.gastro.2015.05.056.

42. Park DJ, Odefrey FA, Hammet F, Giles GG, Baglietto L, Abcfs, Mccs, Hopper JL, Schmidt DF, Makalic E, Sinilnikova OM, Goldgar DE, Southey MC. FAN1 variants identified in multiple-case early-onset breast cancer families via exome sequencing: no evidence for association with risk for breast cancer. Breast cancer research and treatment. 2011; 130: 1043–9. doi: 10.1007/s10549-011-1704-y.

43. Kanchi KL, Johnson KJ, Lu C, McLellan MD, Leiserson MDM, Wendl MC, Zhang Q, Koboldt DC, Xie M, Kandoth C, McMichael JF, Wyczalkowski MA, Larson DE, et al. Integrated analysis of germline and somatic variants in ovarian cancer. Nat Commun. 2014; 5: 3156. doi: 10.1038/ncomms4156.

44. Walsh CS. Two decades beyond BRCA1/2: Homologous recombination, hereditary cancer risk and a target for ovarian cancer therapy. Gynecologic Oncology. 2015; 137: 343–50. doi: 10.1016/j.ygyno.2015.02.017.

45. Mylavarapu S, Das A, Roy M. Role of BRCA Mutations in the Modulation of Response to Platinum Therapy. Front Oncol. 2018; 8: 16. doi: 10.3389/fonc.2018.00016.

46. Stafford JL, Dyson G, Levin NK, Chaudhry S, Rosati R, Kalpage H, Wernette C, Petrucelli N, Simon MS, Tainsky MA. Reanalysis of BRCA1/2 negative high risk ovarian cancer patients reveals novel germline risk loci and insights into missing heritability. Plos One. 2017; 12. doi: 10.1371/journal.pone.0178450.

47. Ceccaldi R, Liu JC, Amunugama R, Hajdu I, Primack B, Petalcorin MIR, O'Connor KW, Konstantinopoulos PA, Elledge SJ, Boulton SJ, Yusufzai T, D'Andrea AD. Homologous-recombination-deficient tumours are dependent on Polθ-mediated repair. Nature. 2015; 518: 258–62. doi: 10.1038/nature14184.

48. Andressoo J-O, Hoeijmakers JHJ, Mitchell JR. Nucleotide excision repair disorders and the balance between cancer and aging. Cell Cycle. 2006; 5: 2886–8. doi: 10.4161/cc.5.24.3565.

7. DISCUSSÃO

A identificação de indivíduos e famílias em risco para o desenvolvimento de câncer hereditário é extremamente importante, uma vez que esses indivíduos apresentam um risco cumulativo vital muito superior ao da população para o desenvolvimento de diversos tipos de câncer. Sendo assim, uma caracterização minuciosa desses indivíduos em risco envolvendo aspectos clínicos, histopatológicos e moleculares poderia auxiliar na identificação, bem como no posterior direcionamento desses indivíduos e seus familiares para programas personalizados de manejo e redução do risco de câncer[17,29]. A HBOC é uma das principais síndromes de predisposição hereditária ao câncer de mama e ovário[86,194]. Porém, os genes até agora associados à HBOC, sejam eles de alto, moderado ou baixo risco, respondem por aproximadamente 50% dos casos, e com isso, as decisões acerca do manejo clínico tornam-se limitadas para os indivíduos em risco para o câncer hereditário sem variantes patogênicas nos genes já associados à HBOC.

No presente estudo, propusemos identificar alterações genéticas em um grupo de mulheres com história pessoal e familiar sugestiva de câncer de mama e/ou ovário hereditário, negativas para variantes patogênicas germinativas nos genes *BRCA1* e *BRCA2*, através de análises em larga escala, tais como sequenciamento exômico e aCGH. Além disso, de forma complementar ao teste genético de *BRCA1*/*BRCA2* já realizado, avaliamos a presença de um rearranjo, comum na população portuguesa, e já classificada como uma variante patogênica, a alteração de c.156_157insAlu no gene *BRCA2*. Para isso, um total de 52 pacientes do sexo feminino, não aparentadas entre si, com história pessoal e familiar de câncer de mama e/ou ovário e alto risco para HBOC foram incluídas. Todas as pacientes foram selecionadas de acordo com a história pessoal e familiar de câncer (amostra por conveniência) e foram previamente encaminhadas para teste genético pelo Departamento de Oncogenética do Hospital de Câncer de Barretos.

Das 52 mulheres analisadas, 41 foram diagnosticadas com câncer de mama, nove com câncer de ovário e duas apresentaram um diagnóstico primário de câncer de ovário, seguido de um segundo tumor primário de câncer de mama. Em relação às características clínico-patológicas das pacientes, a idade média ao diagnóstico de câncer foi de 41,1 anos, variando de 20 a 60 anos. A maioria dos casos de câncer de mama foi do tipo ductal invasivo (85,4%), sendo que segundo a literatura, os tumores "*BRCAness*" também são mais frequentemente

do tipo histológico ductal invasivo[195–197]. Além disso, foi observado em nosso estudo que a maioria das pacientes apresentou positividade para os receptores hormonais (estrógeno: 70,3%, progesterona: 60,5%) e negatividade para HER2 (77,8%). No que se refere ao subtipo molecular, a maioria dos casos de câncer de mama foi classificada molecularmente como tumores do tipo luminal (73,0%), seguido pelos tumores do tipo triplo-negativo (18,9%) e HER2-superexpressos (8,1%). Na literatura já foi demonstrado que a maioria dos tumores "*BRCAness*" exibe um fenótipo triplo-negativo, ou seja, a negatividade para os receptores hormonais e de HER2[198–200]. Ainda, foi observado que a maioria das mulheres com câncer de ovário desenvolveu adenocarcinoma seroso de alto grau (90,9%), sendo que, esse tipo histológico também é mais frequente observado em tumores "*BRCAness*"[201,202].

No que se refere à história familiar das pacientes, além de câncer de mama e ovário, outros cânceres associados ao espectro HBOC, como câncer de próstata (n= 14 familiares, 10 famílias) e pâncreas (n= 4 familiares, 4 famílias) foram observados nas famílias avaliadas no presente estudo. Além disso, casos câncer de mama foram identificados em familiares de primeiro grau em 56% das famílias (n= 29). Ainda, a presença de câncer de ovário entre familiares de primeiro grau foi observada em quatro famílias, e em três dessas quatro famílias o caso índice também foi diagnosticado com esse tipo de tumor. A avaliação da história familiar de câncer é extremamente importante, uma vez que é, reconhecidamente, um fator de risco para o desenvolvimento de síndromes de predisposição hereditária ao câncer[203,204]. Já foi descrito que o risco relativo de câncer de mama é duas vezes maior se houver um familiar de primeiro grau com câncer de mama, e esse risco pode ser até quatro vezes maior se houver dois familiares de primeiro grau com câncer de mama, sendo que o risco aumenta à medida que a idade ao diagnóstico do familiar afetado diminui[205,206]. Recentemente, estudo publicado por Lee *et al.* (2018) relatou que a maioria das mulheres (70,6%) com tumores de mama do tipo "BRCAX" apresentou um diagnóstico de início precoce (≤ 40 anos), seguido de uma história familiar sugestiva de HBOC (42,9%)[207]. Em relação ao nosso estudo, observamos 22 mulheres (42,3%) diagnosticadas com câncer de mama e/ou ovário com idade ≤ 40 anos.

No que se refere às análises moleculares, a avaliação do perfil de alterações cromossômicas por aCGH foi realizada através da plataforma *SureScan Microarray Scanner* (*Agilent Technologies*). Em relação à essa análise, foi realizada uma caracterização tumoral de amostras parafinadas, sendo que um total de 31 amostras foi analisado, após diversos critérios

de qualidade serem empregados e, em virtude desses, nove amostras foram excluídas. Através da análise dos dados, usando o algoritmo GISTIC (*Genomic Identification of Significant Targets in Cancer*), foram identificadas 20 regiões com ganhos genômicos e 31 com perdas.

Dentre as regiões identificadas em nosso estudo, foram observadas distintas alterações no número de cópias envolvendo o cromossomo 8, incluindo perda na região 8p23.3 e ganhos nas regiões 8p12-p11.23 e 8q24.13. Tais resultados apresentaram concordância com estudos prévios envolvendo tumores BRCAX[150,151,208]. Além disso, a região cromossômica 8p12-p11 foi relatada na literatura como amplificada em 10–23% dos casos de câncer de mama[209–211], e ainda, alguns estudos mostraram que a amplificação nessa região está associada a um pior prognóstico[209,212]. Ainda, vale destacar que em no nosso estudo, observamos que das cinco pacientes com o ganho de 8p12-p11, duas (40,0%) apresentaram metástase. Através da realização de análises *in silico* envolvendo a região 8p12-p11.23, na qual identificamos um ganho, destacamos o gene *FGFR1*. Por meio dos dados disponíveis no portal Oncomine[213], pudemos observar uma superexpressão de FGFR1.

O gene *FGFR1* é caracterizado por codificar uma proteína transmembrana que interage com os fatores de crescimento dos fibroblastos e influencia diretamente na diferenciação celular[35]. Além disso, a amplificação do *FGFR1* foi descrita na literatura como a aberração genômica mais frequente em tumores de mama[214,215]. Ainda, diversos estudos já demostraram diferentes resultados no tratamento de mulheres com câncer de mama, dependendo do perfil de *FGFR1*[216–218]. Em 2013, Andrè e colaboradores[219], através da análise de duas linhagens celulares que apresentam amplificação de FGFR1 (MDA-MB-134) e de FGFR2 (SUM52), observaram uma redução na proliferação e no crescimento tumoral após o tratamento com uma droga anti-FGFR1, a Dovitinib (TKI1258). Ainda, os mesmos pesquisadores, demostraram uma diminuição tumoral após o tratamento com Dovitinib em camundongos que sofreram xenoenxerto tumoral de mama com amplificação de *FGFR1*. Adicionalmente, alguns ensaios clínicos foram iniciados e até o presente momento, apenas o Lucitanib (*Clovis Oncology*) apresentou atividade antitumoral significativa (fase I)[220]. De forma complementar, recentemente, o primeiro inibidor de *FGFR* foi aprovado pelo FDA (*U.S. Food and Drug Administration*). O Balversa® (erdafitinib) recebeu aprovação acelerada para o tratamento de pacientes adultos com câncer de bexiga localmente avançado ou metastático

e com alterações genéticas em *FGFR3* ou *FGFR2* e que tenha ocorrido a progressão durante ou após quimioterapia contendo platina.

Adicionalmente, observamos um grande número de amostras com ganhos nas regiões 11q13.2-q13.3 e Xp22.33, as quais foram reportadas anteriormente por Didraga e colaboradores[150]. A superexpressão da cortactina (*CTTN*), presente no 11q13.3, foi associada à amplificação de CCND1 em paciente diagnosticadas com câncer de mama pré-menopausa[221], embora Sheen-Chen *et al.* (2011) não tenha demonstrado um forte valor prognóstico em pacientes com câncer de mama[222]. Embora não seja extensivamente estudado no câncer de mama, a superexpressão de *CRLF2*, presente na região Xp22.33, foi demonstrada como sendo um marcador de pior prognóstico em leucemia linfoblástica aguda pediátrica e adulta do precursor B (conforme publicado por Tasian *et al.*[223]). Em nosso estudo, pudemos observar que das quinze pacientes com um ganho de Xp22.33, três (20%) apresentaram metástase à distância.

Estudo realizado por Gronwald *et al.*[148] comparou os tumores BRCAX com tumores de mama esporádico e identificou diversas regiões alteradas (114 ganhos e 36 perdas) em 18 pacientes. Esses resultados mostraram concordância com resultados obtidos na presente tese, apresentando ganhos mais frequentes em 1q, 6p, 17q e perdas frequentes em 8p. Além dos efeitos bem conhecidos da amplificação do *ERBB2* no desenvolvimento do câncer de mama, a superexpressão do *STARD3* (localizado no mesmo lócus – 17q) parece ser importante, pois pode contribuir para o aumento da proliferação, migração e invasão de células de câncer de mama (conforme publicado por Jacot *et al.* [224]).

Ainda, identificamos que a perda de 22q13.31-13.32 foi associada estatisticamente à presença de câncer de ovário (no probando ou na família). Segundo a literatura, a LOH do cromossomo 22q é relatada em uma variedade de tumores, incluindo os tumores de ovário, em que as taxas de LOH atingiu 70% dos casos[225,226]. Adicionalmente, estudo publicado por Zweemer *et al.* (2001) relatou uma perda significativa de 22q, identificada através de aCGH em tumores de ovário familiar[227]. Curiosamente, através da análise *in silico* envolvendo a região 22q13.31-13.32, identificamos o miR-3201. A expressão do miR-3201 foi reportada na literatura como significativamente reduzido em casos de câncer de ovário epitelial (COE) recorrente, quando comparado aos casos COE primário, indicando possivelmente que se trata de um supressor tumoral[228]. Até onde sabemos, não há estudos que apontam para a

relevância funcional do miR-3201 no câncer de ovário, no entanto, estudos devem ser realizados para avaliar seu possível papel como biomarcador do câncer de ovário.

Em resumo, nossos resultados envolvendo a análise de alterações no número de cópias por aCGH corroboram com os dados reportados anteriormente na literatura envolvendo tumores BRCAX[149–151,208]. Além disso, identificamos novas regiões potencialmente associadas à história pessoal e familiar de câncer de ovário, tais como 22q13.31-13.32. Dada o nosso número amostral limitado, trabalhos adicionais devem ser realizados para validar tais resultados, permitindo assim identificar os genes "*drivers*" associados aos tumores BRCAX no que se refere ao desenvolvimento, bem como para elucidar o papel dessas regiões alteradas na formação e progressão do câncer.

Em relação à análise do rearranjo em *BRCA2* (c.156_157insAlu), foram identificados nove portadores (0,65%) da variante fundadora portuguesa dentre os 1.380 probandos não-relacionados. De acordo com os estudos publicados anteriormente, envolvendo coortes brasileiras e não-brasileiras, esperávamos que aproximadamente 20% de todos os indivíduos que preenchessem os critérios clínicos da NCCN para HBOC apresentassem uma variante patogênica em *BRCA1* ou *BRCA2*[37,229–231]. Portanto, em uma coorte de 1.380 indivíduos, estimamos a presença de aproximadamente 276 portadores de variantes em *BRCA1*/*BRCA2*. Nesse contexto, os nove probandos portadores da variante c.156_157insAlu identificados em nosso estudo representariam cerca de 3,3% de todas as variantes patogênicas detectadas. A menor prevalência mutacional identificada em nosso estudo provavelmente se deve à alta heterogeneidade observada na população brasileira, com contribuições genéticas de diversas regiões do mundo[232]. Essa heterogeneidade da população brasileira é uma consequência de cinco séculos de imigração e cruzamentos interétnicos envolvendo os colonizadores europeus (representados principalmente pelos portugueses), os escravos africanos e os ameríndios. Devido à essa heterogeneidade, realizamos em nosso estudo a análise de ancestralidade, através da utilização de um painel de 46 marcadores (AIM-INDELs). Embora essa análise tenha nos indicado uma maior proporção de ancestralidade Europeia entre todos os portadores da variante fundadora portuguesa (75,3%), um estudo recente mostra que indivíduos das regiões Sudeste e Sul do Brasil derivam de uma ampla região europeia, incluindo a Europa Central e do Norte, bem como partes do Oriente Médio. Esse perfil é diferente dos indivíduos

originários da região Nordeste do Brasil, que podem ter uma ancestralidade europeia restrita principalmente à Península Ibérica[233].

Dentre os nove probandos portadores do rearranjo em *BRCA2*, cinco foram previamente identificados em estudos menores de rastreamento[37,234,235], e a análise de haplótipos foi realizada em três famílias sem, no entanto, uma conclusão definitiva sobre seu haplótipo ancestral[234]. Portanto, no presente estudo, realizamos a reconstrução de haplótipos com nove marcadores de microssatélites. Embora o haplótipo não tenha sido informativo para três famílias, os seis restantes apresentaram haplótipos compatíveis com o haplótipo ancestral português. Assim, esses resultados sugerem que essas famílias compartilham um ancestral comum com as famílias portuguesas positivas para o rearranjo c.156_157insAlu no gene *BRCA2*.

Apesar do grande número de casos brasileiros analisados, nosso estudo é limitado às regiões sudeste e sul do Brasil. Assim, a análise em uma *coorte* maior, incluindo indivíduos de todas as regiões brasileiras, ajudaria a entender a contribuição geral do rearranjo c.156_157insAlu em *BRCA2* na nossa população. Além disso, a prevalência mutacional relativamente baixa encontrada em nosso estudo não descarta a importância de testar essa variante em indivíduos em risco para a HBOC. Vale destacar ainda que embora não seja tão comum quanto em Portugal, a c.156_157insAlu é a terceira variante patogênica em *BRCA2* mais frequente relatada no Brasil, correspondendo a mais de um terço de todos os grandes rearranjos genômicos relatados na população brasileira[38], destacando a importância de sua identificação.

No que se refere à pesquisa de variantes germinativas, através do sequenciamento exômico, um total de 52 mulheres foram incluídas e analisadas (os heredogramas das pacientes incluídas no presente trabalho podem ser encontrados no Anexo VI - LVI). Diversas ferramentas *in silico* foram utilizadas, sendo que, priorizamos os genes já associados à diferentes etapas da carcinogênese, assim como genes de reparo do DNA.

Em relação à análise envolvendo os genes associados à carcinogênese, foram identificadas 53 variantes únicas com perda de função e 128 variantes únicas do tipo *missense*. Além disso, um total de 23 genes que foram descritos como "*Hallmarks of Cancer*" pelo banco de dados COSMIC foram observados[236]. Destes, quatro genes apresentaram variantes que levam à perda de função proteica, e, 19 genes apresentaram variantes *missense* raras (com

frequência populacional ≤0.02 pelo banco de dados gnomAD[237]). Alguns desses genes foram destacados em nosso estudo devido à sua função e à possível associação com tumores hereditários/familiares.

O gene mais frequentemente alterado foi o *CCND3*, sendo que foi identificada uma variante *nonsense* (c.379G>T) em quatro mulheres não relacionadas (ID dos pacientes: 80, 85, 179 e 1014). Essas pacientes foram diagnosticadas com câncer de mama, apresentando uma idade média ao diagnóstico de 46 anos. Ainda, todas as portadoras de variantes em *CCND3* relataram história familiar de câncer de mama. A presença dessa variante foi confirmada por sequenciamento de Sanger nas quatro pacientes. O gene *CCND3* codifica uma proteína que funciona na regulação de quinases dependentes de ciclina no ciclo celular[238]. Um estudo independente relatou a presença da mesma variante (c.379G>T) identificada em nosso estudo, em tecido germinativo, em 1,61% (9/557) dos pacientes caucasianos diagnosticados com câncer de ovário[188]. Segundo as diretrizes do ACMG-AMP, essa variante é classificada como de significado incerto, e ainda, não foi reportada pelo ClinVar. Ao nosso conhecimento, variantes germinativas em *CCND3* não foram reportadas em pacientes com câncer de mama.

A variante *nonsense* identificada no gene *DROSHA* (c.1498G>T), não reportada pelo ClinVar e classificada como patogênica pelas diretrizes do ACMG-AMP, foi identificada em um paciente (ID 133) com história pessoal e familiar de câncer de mama. Embora essa variante não tenha sido reportada pelo ClinVar ou na literatura, esse gene foi descrito como crucial na biogênese de microRNAs e na interação direta com efetores de p53 associados à ligação de RNA[238]. Estudo realizado por Qian e colaboradores identificou um polimorfismo em *DROSHA* (rs78393591) em mulheres de ascendência africana diagnosticadas com câncer de mama[239]. Além disso, variantes somáticas envolvendo esse gene têm se mostrado altamente frequente no tumor de Wilms[240]. Portanto, essas observações sugerem um possível papel de *DROSHA* na etiologia do câncer.

Outro gene descrito como "*Hallmarks of Cancer*", o *SLC34A2*, foi identificado em nosso estudo com uma variante afetando o *splicing* (c.113-2A>G). A variante foi identificada em uma paciente com história pessoal e familiar de câncer de ovário. Adicionalmente, essa variante não foi reportada pelo ClinVar e foi classificada como uma variante de significado incerto pelas diretrizes da ACMG-AMP. Estudo realizado por Kanchi *et al.* (2014), através da análise de variantes germinativas e somáticas em câncer de ovário, observou a presença de uma variante

germinativa em um sítio de *splicing* (c.1458+2T>C) e uma variante *missense* (c.1079C>T)[188] envolvendo esse gene considerado um supressor tumoral. No entanto, como essas variantes foram encontradas em baixa frequência, os autores não puderam associar variantes em *SLC34A2* ao desenvolvimento do câncer de ovário.

Ainda, identificamos variantes germinativas nos genes associados ao sistema de reparo de DNA *mismatch repair* (MMR). Dentre os genes alterados, destacam-se o *MLH1*, *MSH2*, *PMS2* e *MSH6*, os quais são classicamente associados à síndrome de Lynch[241,242]. No presente estudo, identificamos variantes nos genes *MLH1* (c.794G>A), *MSH2* (c.80C>T), *MSH6* (c.2885T>C) e *PMS2* (c.2182_2184delinsG) em pacientes com história pessoal e familiar de câncer de mama. Dentre as variantes nos genes do MMR, destacamos a variante no gene *PMS2* (c.2182_2184delACTinsG), por se tratar de uma alteração do tipo *frameshift* reportada como patogênica pelo banco de dados ClinVar e pelos critérios da ACMG-AMP. Tal variante (c.2182_2184delACTinsG) foi observada em uma mulher (ID 306) diagnosticada com melanoma (aos 26 anos) seguido por diagnóstico de câncer de mama (aos 36 anos). Adicionalmente, através da análise de segregação, observamos a presença da mesma variante no gene *PMS2* (c.2182_2184delACTinsG) em três membros de primeiro grau testados (irmãs), sendo que, apenas um membro foi diagnosticado com dois tumores primários: câncer de mama (aos 42 anos de idade) seguido de câncer de tireoide (aos 43 anos de idade). Alguns pesquisadores já descreveram uma associação de variantes patogênicas em genes do MMR com câncer de mama, embora uma discrepância seja observada na literatura. Um estudo realizado por Couch *et al.* (2017), analisou 65.057 mulheres não-portadoras de variantes patogênicas em *BRCA1* e *BRCA2* e diagnosticadas com câncer de mama. Os autores estimaram um risco baixo/moderado para o desenvolvimento de câncer de mama em portadores de variantes patogênicas, sendo esses riscos de 1,15 para *MLH1* (IC 95%: 0,30-4,19), 2,46 para *MSH2* (IC 95%: 0,81-6,93), 1,93 para *MSH6* (IC 95%: 1,16-3,27) 0,82 para *PMS2* (IC 95%: 0,44-1,47)[86]. Além disso, segundo as diretrizes do NCCN os genes classicamente associados à Síndrome de Lynch não apresentam evidências suficientes de associação com um risco de desenvolvimento de câncer de mama[42]. No entanto, Roberts *et al.* (2018), através de um estudo caso-controle, relatou uma associação entre variantes patogênicas em *PMS2* e um aumento no risco para o desenvolvimento de câncer de mama, em mulheres testadas negativas para *BRCA1*/*BRCA2*. Devido à alta frequência (29,3%) de variantes germinativas em

PMS2 (124/423), os autores concluíram que esse gene pode estar associado a um risco aumentado de câncer de mama e deve ser considerado em mulheres com história pessoal e familiar de câncer de mama[182]. Além disso, LaDuca *et al.* (2019), observou que 37,4% dos pacientes com variantes patogênicas em *PMS2* apresentaram critérios clínicos para a indicação de teste genético para *BRCA1*/*BRCA2*. Assim, os autores destacaram a importância da implementação de painéis gênicos em pacientes em risco para o câncer hereditário, mesmo quando os pacientes apresentam características clínicas clássicas de uma determinada síndrome de predisposição ao câncer, resultando em oportunidades de tratamento e acompanhamento para as famílias[113].

Ainda, neste estudo, identificamos outro gene previamente associado ao câncer colorretal, o *POLQ*[243–245]. Curiosamente, a variante *frameshift* identificada em *POLQ* (c.4262_4268del) foi observada em duas pacientes não-relacionadas e diagnosticadas com câncer de mama, aos 42 e 47 anos, respectivamente. Estudos publicados por Wang *et al.* (2008)[244] e Brandalize *et al.* (2014)[245] relataram a presença de variantes germinativas em *POLQ* em pacientes com história pessoal e familiar de câncer de mama e não-portadoras de variantes patogênicas em *BRCA1*/*BRCA2*. Além disso, um estudo caso-controle publicado por Family *et al.* (2015)[246], associou três variantes *missense* em *POLQ* com um risco aumentado para o desenvolvimento de câncer de mama. Assim, os autores concluíram que *POLQ* poderia ser considerado um possível gene candidato envolvido no desenvolvimento do câncer de mama hereditário.

Variantes patogênicas/provavelmente-patogênicas foram identificadas em *RAD51C* e *CHEK2*. Esses genes já foram associados à um risco aumentado para o desenvolvimento de câncer de mama e/ou ovário e participam da via de reparo por recombinação homóloga[86]. A variante identificada em *RAD51C* (c.890_899delTTGTTCCTGC) foi observada em uma paciente (ID 656) diagnosticada com câncer de ovário (aos 41 anos) e câncer de mama (aos 48 anos), sendo que na história familiar de câncer foi reportado a presença de câncer de ovário em um familiar de segundo grau. Alguns estudos mostram que *RAD51C* pode representar um gene que confere um risco moderado a alto para o desenvolvimento de câncer de ovário (OR: 4,98 [IC 95%: 3,09-8,04][62] e OR: 5,2 [IC 95%: 1,1-24,0][102]). Além disso, de acordo com as diretrizes da NCCN, o risco cumulativo vital para o desenvolvimento de câncer de ovário nos portadores de variantes patogênicas em *RAD51C* parece ser suficiente para justificar a consideração de

salpingo-ooforectomia para redução de risco[42]. Em relação ao câncer de mama, estudos mostraram que a presença de variantes patogênicas em *RAD51C* não foi relacionado à um aumento no riscos (OR: 0,78 [IC 95%: 0,47-1,37][86] e 1,43 [IC 95%: 0,97-2,12][62]). No entanto, Shimelis *et al*. (2018) reportaram que variantes patogênicas em *RAD51C* foram associadas a risco moderado (OR: 2,64; IC 95%: 1,44-4,80) para o desenvolvimento de câncer de mama triplo-negativo[60]. Já a variante patogênica identificada no gene *CHEK2* (c.349A> G) foi observada em uma mulher (ID 1326) com história pessoal e familiar de câncer de mama (diagnóstico aos 35 anos). O gene *CHEK2* já foi associado a um risco aumentado de câncer de mama (OR: 2,26 [IC 95%: 1,89-2,72])[86]. Um estudo recente publicado por Kleiblova *et al.* (2019), através da análise de *CHEK2* em 1.928 pacientes em alto risco para a HBOC, identificou 10 variantes que levam à perda de função proteica em 46 (2,39%) pacientes e 26 variantes do tipo *missense* em 88 (4,56%) pacientes. Ainda, através da análise *in silico*, os autores inferiram que a presença de variantes raras e danosas em *CHEK2* apresentava um risco significativamente aumentado para o desenvolvimento de câncer de mama (OR: 3,90; IC 95%: 1,24-13,35) e câncer de ovário (OR: 4,77; IC 95%: 0,77-22,47)[178]. Além disso, estudo publicado por Southey *et al.* (2016) objetivou estimar o risco de desenvolvimento de câncer associado à algumas variantes, dentre elas as variantes identificadas em nosso estudo (c.349A>G). Os autores estimaram que a variante c.349A>G apresentou um risco de 2,26 (IC 95%: 1,29-3,95), indicando um risco moderado para o desenvolvimento de câncer de mama[179]. De acordo com as diretrizes do NCCN, a mamografia anual é indicada para os portadores de variantes patogênicas em *CHEK2*[42].

É importante enfatizar que identificamos algumas variantes do tipo *missense* em oncogenes, como *EGFR*, *ERBB4*, *KIT* e *KRAS*. Esses genes são descritos como superfamília-quinase[238], sendo que a amplificação envolvendo esses genes já foram associadas à diversos tipos de tumores, como tumores de mama[247], gastrointestinal[248,249] e glioblastomas[250,251]. Vale destacar que alguns estudos reportaram a presença de variantes germinativas envolvendo alguns desses oncogenes em tumores de mama/ovário familiar. Penkert *et al.* (2018), realizou o sequenciamento de 94 genes de predisposição para o câncer em uma *coorte* alemã de 83 mulheres diagnosticadas com câncer de mama e sem variantes patogênicas germinativas em *BRCA1*, *BRCA2* e *TP53*. Os autores relataram a presença de variantes germinativas do tipo *missense* em *EGFR* (c.2039G>A) e *KIT* (c.391G>A), ambos com frequência de 1,2%[252]. Em nossa

coorte, as variantes identificadas nesses genes foram observadas em pacientes com câncer de ovário (*EGFR*: c.352G>T, ID 1231), com câncer de mama (*KIT:* c.50T>C, ID: 426) e câncer de mama e ovário (*ERBB4*: c.3446G>T e *KRAS*: c.461A>G, ambas variantes indentificadas na paciente ID 29). As variantes identificadas em *KRAS* e *EGFR* estavam localizadas em domínios que ativam esses proto-oncogenes (domínios Ras e Receptor-L, respectivamente)[253(p2)]. É importante ressaltar que a associação das variantes identificadas nesses oncogenes será posteriormente avaliada por análise de co-segregação, uma vez que essas variantes foram confirmadas por sequenciamento de Sanger.

Em relação as vias de reparo à danos ao DNA, das 52 mulheres incluídas e com história pessoal e familiar sugestiva de câncer de mama/ovário hereditário, 29 (55,8%) pacientes apresentaram variantes germinativas em 26 diferentes genes de reparo ao DNA. Identificamos um total de 35 variantes, sendo nove variantes que levam à perda de função proteica e 26 do tipo *missense*. No que se refere à classificação de patogenicidade segundo a ACMG-AMP, 29 (82,9%) variantes foram classificados como de significado incerto, seguidos por quatro (11,4%) provavelmente-patogênicas/patogênicas e duas (5,7%) provavelmente-benignas/benignas.

As vias de reparo por Recombinação Homóloga, *Mismatch Repair* e Anemia de Fanconi além de apresentaram uma alta frequência de variantes, identificamos variantes potencialmente patogênicas envolvendo os genes *RAD51C, CHEK2*, *PMS2* (discutidas nos parágrafos anteriores) e *FAN1*. Adicionalmente, destacamos duas variantes do tipo *missense* envolvendo o gene *RAD54L*, discutidas abaixo.

Quanto à via de Anemia de Fanconi, destacamos o gene *FAN1*, o qual foi observado a presença de variantes germinativas em quatro pacientes, sendo que todos os casos apresentaram história pessoal e familiar de câncer de mama. Vale destacar que, apenas uma variante identificada em nosso estudo foi caracterizada por levar à perda de função proteica (c.357_358delGG, identificada na paciente ID 974). Tal variante foi descrita como provável patogênica pelo ACMG-AMP e não descrita pelo ClinVar. Adicionalmente, a variante c.149T>G, não foi reportada pelo ClinVar e considerada como de significado incerto pelo ACMG-AMP, foi identificada em duas pacientes (ID 426 e 1264). Essa mesma variante foi reportada anteriormente por Smith *et al.* (2016) em pacientes com história pessoal e familiar de câncer pancreático. Ainda, os autores observaram a co-segregação da variante em duas famílias diferentes, considerando assim, um gene candidato à predisposição ao câncer pancreático[183].

Diante do exposto, Lachaud *et al.* (2016) por meio de ensaios *in vitro*, objetivou avaliar o impacto da variante c.149T>G. Os autores reportaram que, devido à localização da variante em um domínio funcional (UBZ: domínio de ligação à ubiquitina), houve a perda da função proteica. Portanto, os pesquisadores concluíram que a variante c.149T>G foi associada ao desenvolvimento de câncer de pâncreas[184]. Adicionalmente, alguns autores têm relatado a possível associação de variantes germinativas em *FAN1* com o câncer colorretal[185,186]. No entanto, poucos estudos relataram a possível associação da presença de variantes germinativas em *FAN1* e o desenvolvimento de câncer de mama. Park *et al*. (2011)[187], através do sequenciamento exômico de famílias diagnosticadas com câncer de mama de início precoce, provenientes da Austrália, identificou a presença de variantes no gene *FAN1*. Através da análise de segregação, os autores observaram que a variante c.1129C>T foi identificada em todos os seis familiares testados e diagnosticados com câncer (cinco com câncer de mama e um com melanoma). Portanto, estudos futuros precisam ser realizados para melhor elucidar o possível papel do gene *FAN1* associado ao desenvolvimento do câncer de mama.

No que se refere às variantes identificadas no gene *RAD54L* (c.604C>T e c.1094G>A), ambas foram classificadas como de significado clínico desconhecido pelas diretrizes da ACMG-AMP e não reportadas pelo ClinVar. Adicionalmente, observamos que as duas variantes foram identificadas em pacientes com história pessoal e familiar de câncer de ovário (IDs 320: c.604C>T e 565: c.1094G>A). Embora a função do *RAD54L* na síndrome HBOC ainda não esteja clara, alguns autores relataram a presença de variantes germinativas em pacientes diagnosticadas com câncer de mama e/ou ovário. Matsuda *et al.* (1999)[180] identificou a presença de uma variante somática (c.973G> A) em uma mulher diagnosticada com câncer de mama aos 63 anos de idade e sem história familiar de câncer. Ainda, os autores observaram a ausência do alelo *wildtype* no tecido normal, indicando se tratar de uma variante germinativa com características de um supressor tumoral. Além disso, os autores reportaram a ausência dessa variante em 100 indivíduos “controles” (sem o diagnóstico de câncer). Shagimardanova *et al*. (2018)[181], através de dados publicados no *Annals of Oncology*, analisou 568 pacientes diagnosticados com câncer de mama e/ou ovário, entre 21 e 82 anos de idade. Os autores relataram que 193/568 (34%) pacientes apresentaram características clínicas da síndrome HBOC. Ainda, 128 (22,5%) pacientes foram portadores de variantes patogênicas ou prováveis patogênicas em *BRCA1*; 52 (9,2%) em *BRCA2* e 100 (17,6%) em um dos outros genes

relacionados ao HBOC, incluindo o *RAD54L*. Maiores detalhes sobre as variantes em genes não-*BRCA1*/*BRCA2* não foram publicados pelos autores.

Ainda, enfatizamos que variante em genes envolvidos na via de Reparo por Excisão de Base (BER- Base Excision Repair) foram identificados em nosso estudo. Foram observadas três variantes que levam à perda de função proteica envolvendo os genes *NEIL1* (c.434+2T>C), *POLQ* (c.4262_4268del) e *TDG* (c.1090_1090+1insTTGAGAGC) e quatro variantes do tipo *missense* nos genes *FEN1* (c.311G>A), *RECQL* (c.401C>T) e *UNG* (c.262C>T) classificadas como de significado incerta de acordo com os critérios AMCG-AMP. Poucos estudos relataram a presença de variantes germinativas nesses genes em pacientes de alto risco para câncer de mama/ovário hereditário. No entanto, Kanchi *et al.* (2014) relataram a presença de variantes que levam à perda de função proteica envolvendo os genes *NEIL1* e *UNG* em mulheres diagnosticadas com câncer de ovário[188]. Além disso, alguns autores têm proposto que a presença de variantes patogênicas em genes das vias de Reparo por Excisão de Base e de Recombinação Homóloga levaria à morte[189,190] e, com isso, os pacientes poderiam se beneficiar de terapias com inibidores de PARP. Diante do exposto, em nossa *coorte* observamos dois pacientes com variantes envolvidas nas vias duas vias (reparo por excisão de base e recombinação homóloga). A paciente ID 565 apresentou história pessoal e familiar de câncer de ovário, e foi identificada com as variantes c.434+2T>C (no gene *NEIL1*) e c.604C>T (no gene *RAD54L*). Já a paciente ID 1097, diagnosticada com câncer de ovário e com história familiar de câncer de mama, foi observada com uma variante que leva à perda de função proteica em *TDG* (c.1090_1090+1insTTGAGAGC) e uma variante *missense* em *RAD50* (c.353T>C).

Não menos importante, encontramos variantes germinativas na via de Reparo por Excisão de Nucleotídeos (NER, *Nucleotide Excision Repair*). Dentre as variantes identificadas, ambas classificadas como variante de significado clínico desconhecido pelo ACMG-AMP, uma foi observada no gene *ERCC6* (c.1801G>A) e outra no gene *ERCC8* (c.839C>A). A variante identificada em *ERCC6* foi observada em uma paciente (ID 1231) diagnosticada com câncer de ovário, enquanto a variante identificada em *ERCC8* foi observada em uma paciente (ID 179) diagnosticada com câncer de mama, sendo que ambas as pacientes reportaram a presença de câncer de mama na história familiar. Variantes germinativas nos genes da família *ERCC* têm sido associadas ao Xeroderma Pigmentoso, uma síndrome que aumenta o risco de vários tipos

de câncer[193]. Através da análise de tecido tumoral de câncer de ovário, Kanchi *et al.* (2014) identificou uma variante que leva à produção de uma proteína truncada em *ERCC8*[188]. No entanto, a presença de variantes germinativa envolvendo os genes *ERCC6* e *ERCC8* não foi reportada na literatura (ao nosso conhecimento) e uma possível associação com o desenvolvimento de câncer de mama/ovário nos casos diagnosticados em idade precoce não é clara.

Embora o sequenciamento exômico tenha se tornado uma abordagem eficaz para identificar variantes em alelos e, assim, novos genes candidatos à predisposição genética em diversas doenças hereditárias, essa técnica apresenta inúmeros desafios, tais como as diferentes ferramentas de bioinformática utilizadas para avaliação dos dados, bem como a priorização das variantes e a variabilidade específica da população. Ainda, o sequenciamento exômico realizado em nosso estudo, não nos permite analisar as regiões intrônicas, bem como grandes rearranjos, mutações do genoma mitocondrial e mutações em mosaico, representando assim uma limitação do presente estudo.

Por fim, realizamos uma integração dos dados provenientes das duas tecnologias utilizadas na presente tese. Para isso, as regiões identificadas com ganhos e perdas genômicas (pelo aCGH) no tecido tumoral das pacientes foram comparadas à localização da citobanda de cada gene e com as variantes identificadas nesses genes (pelo sequenciamento do exoma) no DNA constitutivo. Através dessa análise, identificamos duas regiões que apresentaram alterações tanto na linhagem germinativa quanto no tumor. Destas regiões, destacamos a perda de 4q35.2 e 5p15.33, discutidas a seguir.

No que se refere à região cromossômica 4q35.2 podemos destacar a presença do gene *FAT1*, considerado um supressor tumoral ou oncogene dependente do contexto[254,255], descrito pelo banco de dados COSMIC como "*Hallmarker of Cancer*"[236]. Foi identificado uma variante do tipo *missense* (c.9583T>A), a qual apresentou uma frequência alélica ≤ 0,02 nos bancos de dados populacionais, gnomAD[237] e AbraOM[169]. Além disso, essa variante foi predita como provavelmente-patogênica pelas ferramentas *in silico* REVEL (*score*: 0,805), CADD (*score*: 28,6) e M-CAP (*score*: 0,14). Vale destacar que a variante c.9583T>A foi observada em uma paciente (ID 960) diagnosticada com câncer de mama bilateral (aos 59 e 70 anos, respectivamente) e ainda, reportou na história familiar a presença de câncer de mama (duas familiares, um de primeiro grau e o outro de terceiro grau). Variantes germinativas em *FAT1*

foram reportadas na literatura em alguns tumores, como pâncreas[256], colorretal[257] e gástrico[258]. Adicionalmente, a perda da região 4q35.2 foi reportada por Johannsdottir e colaboradores (2004) em pacientes diagnosticados com câncer de mama. Os autores observaram a perda envolvendo essa região (4q35.2) em 11/15 (73%) mulheres com variantes germinativas em *BRCA1*, 14/38 (37%) mulheres com variantes germinativas em *BRCA2* e 7/36 (19%) de mulheres não-portadoras de variantes germinativas em *BRCA1*/*BRCA2*[259].

Já em relação à outra região que identificamos perda, a 5p15.33, podemos destacar a presença do gene *TERT*. Segundo o banco de dados COSMIC[236], o gene *TERT* desempenha tanto as funções associadas à supressão tumoral quanto funções de promoção tumoral (oncogene), sendo considerado um "*Hallmarker of Cancer*". Sobre a variante identificada em *TERT*, é uma alteração do tipo *missense* (c.3332C>T), classificada como provavelmente patogênica pelas ferramentas *in silico* (REVEL: *score* de 0,56; CADD: *score* de 24,3 e M-CAP: *score* de 0,80). A paciente (ID 275) que apresentou essa variante em *TERT* foi diagnosticada com câncer de ovário aos 60 anos de idade, e, ainda, reportou a presença de dois casos de câncer de mama na história familiar (ambos familiares de segundo grau). Variantes germinativas na região promotora de *TERT* foram relatadas inicialmente por Horn *et al.* (2013)[260] em pacientes com melanoma familial. Os autores observaram que os portadores de variantes germinativas em *TERT* apresentaram um início precoce ao diagnóstico com uma rápida progressão e o desenvolvimento de diversos tumores. Vale enfatizar ainda que variantes germinativas bialélicas em *TERT* também já foram associadas à disceratose congênita, uma síndrome rara caracterizada por falência da medula óssea, anormalidades mucocutâneas e predisposição a alguns tipos de câncer, como colorretal, pâncreas e pulmão[261].

De maneira geral, os estudos descritos acima bem como os resultados obtidos até o momento evidenciam a existência de alterações germinativas e somáticas em famílias com ausência de variantes patogênicas em *BRCA1*/*BRCA2*, e que por sua vez podem contribuir para o desenvolvimento do câncer de mama e/ou ovário hereditário. O objetivo geral deste estudo foi avaliar mulheres que foram diagnosticadas com câncer de mama e/ou ovário, com história familiar destes tumores. Os critérios de inclusão e exclusão foram determinados levando em conta as características clínicas e patológicas das pacientes, bem como a avaliação acurada da história familiar com a finalidade de reduzir o viés de uma casuística muito complexa. Em

suma, os genes identificados e destacados em nosso estudo parecem interagir de alguma maneira, ou por desempenharem um papel importante no processo da carcinogênese ou em vias de reparo. As técnicas aqui empregadas possibilitaram a identificação de alterações gênicas e cromossomais. Através deste estudo identificamos novas alterações germinativas e somáticas em pacientes com câncer de mama e/ou ovário, além de alterações já reportadas na literatura, revelando assim potenciais candidatos à predisposição hereditária para o desenvolvimento destes tumores. É importante ainda ressaltar a necessidade de realizar análises de co-segregação e pesquisa da perda de heterozigose (LOH). As mesmas serão realizadas e analisadas futuramente, a fim de fornecer maiores evidências acerca do impacto clínico e biológico da presença das variantes germinativas aqui identificadas.

8. CONCLUSÕES

As principais observações e conclusões destas investigações podem ser resumidas da seguinte forma:

- A frequência da inserção Alu em *BRCA2* no éxon 3 (c.156_157insAlu) entre indivíduos que preencheram os critérios da síndrome de Câncer de Mama e Ovário Hereditário (HBOC) foi de 0,65% (9/1.380 casos avaliados). Nos probandos portadores da variante fundadora portuguesa, a ancestralidade europeia apresentou a maior proporção (80%), seguida pela africana (10%) e ameríndia. Na maioria das famílias com o rearranjo, as análises de haplótipos indicaram presença do haplótipo ancestral português.

- Através da análise de aCGH, foram identificadas 20 regiões com ganhos e 31 regiões com perdas. Dentre os cromossomos com mais regiões alteradas, destacamos os cromossomos 8 (em tumores de mama e ovário) e X (em tumores de ovário). Através da análise *in silico* envolvendo a região 8p12-p11.23, identificamos o gene *FGFR1*. Além disso, identificamos que a perda de 22q13.31-13.32 foi associada estatisticamente à presença de câncer de ovário (no probando ou na família). Nessa região observamos, através da análise *in silico,* a presença do miR-3201. Ainda, o ganho de 6p22.1 e perda de 6q25.1 foram associados à presença de metástase.

- No presente estudo, através do sequenciamento exômico, foi identificado quatro variantes patogênicas/provavelmente-patogênicas pelo ClinVar/ACMG-AMP. Dentre elas, destacam-se os genes *CHEK2*, *FAN1*, *RAD51C*, *PMS2*. Além disso, as vias de reparo por Recombinação Homóloga, *Mismatch Repair* e Anemia de Fanconi foram as mais frequentemente alteradas. O presente estudo fornece evidências adicionais para a associação de genes considerados de moderado risco para o desenvolvimento de câncer de mama e/ou ovário.

- Os resultados do presente estudo contribuirão para a catalogação molecular de tumores de mama e/ou ovário em pacientes em risco para HBOC, previamente

testados negativos para variantes germinativas patogênicas em *BRCA1*/*BRCA2*, bem como o desenvolvimento de estratégias de detecção precoce baseadas em genes e terapias direcionadas como foi estabelecido em portadores de variantes patogênicas em *BRCA1*/*BRCA2*.

Referências

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*. 2018;68(6):394-424. doi:10.3322/caac.21492

2. UICC. Introduction UICC Global Cancer Control.

3. INCA. Estimativa 2018: Incidência de Câncer no Brasil. 2018.

4. Yokota J. Tumor progression and metastasis. *Carcinogenesis*. 2000;21:497-503.

5. Renan MJ. How many mutations are required for tumorigenesis? Implications from human cancer data. *Molecular carcinogenesis*. 1993;7:139-146.

6. Weinberg RA. Tumor suppressor genes. *Science*. 1991;254:1138-1146.

7. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144:646-674. doi:10.1016/j.cell.2011.02.013

8. Beckmann MW, Niederacher D, Schnurch HG, Gusterson BA, Bender HG. Multistep carcinogenesis of breast cancer and tumour heterogeneity. *Journal of molecular medicine*. 1997;75:429-439.

9. Balmain A, Gray J, Ponder B. The genetics and genomics of cancer. *Nature genetics*. 2003;33 Suppl:238-244. doi:10.1038/ng1107

10. Weinberg RA. Oncogenes, antioncogenes, and the molecular bases of multistep carcinogenesis. *Cancer research*. 1989;49:3713-3721.

11. Osborne C, Wilson P, Tripathy D. Oncogenes and tumor suppressor genes in breast cancer: potential diagnostic and therapeutic applications. *The oncologist*. 2004;9:361-377.

12. Knudson AG. Mutation and cancer: statistical study of retinoblastoma. *Proceedings of the National Academy of Sciences of the United States of America*. 1971;68:820-823.

13. Sabado Alvarez C. Molecular biology of retinoblastoma. *Clinical & translational oncology : official publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico*. 2008;10:389-394.

14. Miller BJ, Wang D, Krahe R, Wright FA. Pooled analysis of loss of heterozygosity in breast cancer: a genome scan provides comparative evidence for multiple tumor suppressors and identifies novel candidate regions. *American journal of human genetics*. 2003;73:748-767. doi:10.1086/378522

15. Easton DF, Pharoah PD, Antoniou AC, et al. Gene-panel sequencing and the prediction of breast-cancer risk. *The New England journal of medicine*. 2015;372:2243-2257. doi:10.1056/NEJMsr1501341

16. Samadder NJ, Giridhar KV, Baffy N, Riegert-Johnson D, Couch FJ. Hereditary Cancer Syndromes-A Primer on Diagnosis and Management: Part 1: Breast-Ovarian Cancer Syndromes. *Mayo Clin Proc*. 2019;94(6):1084-1098. doi:10.1016/j.mayocp.2019.02.017

17. Melchor L, Benítez J. The complex genetic landscape of familial breast cancer. *Human Genetics*. 2013;132(8):845-863. doi:10.1007/s00439-013-1299-y

18. Walsh T, Casadei S, Lee MK, et al. Mutations in 12 genes for inherited ovarian, fallopian tube, and peritoneal carcinoma identified by massively parallel sequencing. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;108:18032-18037. doi:10.1073/pnas.1115052108

19. Jones MR, Kamara D, Karlan BY, Pharoah PDP, Gayther SA. Genetic epidemiology of ovarian cancer and prospects for polygenic risk prediction. *Gynecologic Oncology*. 2017;147(3):705-713. doi:10.1016/j.ygyno.2017.10.001

20. Krush AJ. *Contributions of Pierre Paul Broca to Cancer Genetics*. Vol VII. University of Nebraska - Lincoln; 1979.

21. Anderson DE. Genetic study of breast cancer: identification of a high risk group. *Cancer*. 1974;34:1090-1097.

22. Hauser AR, Lerner IJ, King RA. Familial male breast cancer. *American journal of medical genetics*. 1992;44:839-840. doi:10.1002/ajmg.1320440626

23. Houlston RS, McCarter E, Parbhoo S, Scurr JH, Slack J. Family history and risk of breast cancer. *Journal of medical genetics*. 1992;29:154-157.

24. Petrakis NL. Genetic factors in the etiology of breast cancer. *Cancer*. 1977;39:2709-2715.

25. Wooster R, Bignell G, Lancaster J, et al. Identification of the breast cancer susceptibility gene BRCA2. *Nature*. 1995;378:789-792. doi:10.1038/378789a0

26. Miki Y, Swensen J, Shattuck-Eidens D, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science*. 1994;266:66-71.

27. de Silva D, Gilbert F, Needham G, Deans H, Turnpenny P, Haites N. Identification of women at high genetic risk of breast cancer through the National Health Service Breast Screening Programme (NHSBSP). *Journal of medical genetics*. 1995;32:862-866.

28. Felix GES, Zheng Y, Olopade OI. Mutations in context: implications of BRCA testing in diverse populations. *Familial Cancer*. 2018;17(4):471-483. doi:10.1007/s10689-017-0038-2

29. Couch FJ, Nathanson KL, Offit K. Two decades after BRCA: setting paradigms in personalized cancer care and prevention. *Science*. 2014;343:1466-1470. doi:10.1126/science.1251827

30. Mehta A. BRCA1 and BRCA2 mutations in ovarian cancer. *Journal of Current Oncology*. 2018;1(1):1. doi:10.4103/jco.jco_6_18

31. Mahdavi M, Nassiri M, Kooshyar MM, et al. Hereditary breast cancer; Genetic penetrance and current status with BRCA. *Journal of Cellular Physiology*. December 2018. doi:10.1002/jcp.27464

32. Michailidou K, Lindström S, Dennis J, et al. Association analysis identifies 65 new breast cancer risk loci. *Nature*. 2017;551(7678):92-94. doi:10.1038/nature24284

33. Sawyer S, Mitchell G, McKinley J, et al. A role for common genomic variants in the assessment of familial breast cancer. *J Clin Oncol*. 2012;30:4330-4336. doi:10.1200/JCO.2012.41.7469 JCO.2012.41.7469 [pii]

34. Foulkes WD. Inherited Susceptibility to Common Cancers. *New England Journal of Medicine*. 2008;359(20):2143-2153. doi:10.1056/NEJMra0802968

35. The UniProt Consortium. UniProt: the universal protein knowledgebase. *Nucleic Acids Res*. 2017;45(D1):D158-D169. doi:10.1093/nar/gkw1099

36. El Khachibi M, Diakite B, Hamzi K, et al. Screening of exon 11 of BRCA1 gene using the high resolution melting approach for diagnosis in Moroccan breast cancer patients. *BMC cancer*. 2015;15:81. doi:10.1186/s12885-015-1040-4

37. Fernandes GC, Michelli RA, Galvao HC, et al. Prevalence of BRCA1/BRCA2 mutations in a Brazilian population sample at-risk for hereditary breast cancer and characterization of its genetic ancestry. *Oncotarget*. 2016;7:80465-80481. doi:10.18632/oncotarget.12610

38. Palmero EI, Carraro DM, Alemar B, et al. The germline mutational landscape of BRCA1 and BRCA2 in Brazil. *Sci Rep*. 2018;8(1):9188. doi:10.1038/s41598-018-27315-2

39. Rebbeck TR, Friebel TM, Friedman E, et al. Mutational spectrum in a worldwide study of 29,700 families with BRCA1 or BRCA2 mutations. *Hum Mutat*. 2018;39(5):593-620. doi:10.1002/humu.23406

40. Kuchenbaecker KB, Hopper JL, Barnes DR, et al. Risks of Breast, Ovarian, and Contralateral Breast Cancer for *BRCA1* and *BRCA2* Mutation Carriers. *JAMA*. 2017;317(23):2402. doi:10.1001/jama.2017.7112

41. Fostira F, Tsitlaidou M, Papadimitriou C, et al. Prevalence of BRCA1 mutations among 403 women with triple-negative breast cancer: implications for genetic screening selection criteria: a Hellenic Cooperative Oncology Group Study. *Breast cancer research and treatment*. 2012;134:353-362. doi:10.1007/s10549-012-2021-9

42. National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines) for Genetic/Familial High-Risk Assessment - v3.2019. https://www.nccn.org/professionals/physician_gls/pdf/genetics_screening.pdf.

43. Castro E, Eeles R. The role of BRCA1 and BRCA2 in prostate cancer. *Asian J Androl*. 2012;14(3):409-414. doi:10.1038/aja.2011.150

44. Ibrahim M, Yadav S, Ogunleye F, Zakalik D. Male BRCA mutation carriers: clinical characteristics and cancer spectrum. *BMC Cancer*. 2018;18(1):179. doi:10.1186/s12885-018-4098-y

45. Greer JB, Whitcomb DC. Role of BRCA1 and BRCA2 mutations in pancreatic cancer. *Gut*. 2007;56(5):601-605. doi:10.1136/gut.2006.101220

46. Antoniou AC, Pharoah PDP, Easton DF, Evans DG. BRCA1 and BRCA2 cancer risks. *J Clin Oncol*. 2006;24(20):3312-3313; author reply 3313-3314. doi:10.1200/JCO.2006.06.7934

47. Ha SM, Chae EY, Cha JH, Kim HH, Shin HJ, Choi WJ. Association of BRCA Mutation Types, Imaging Features, and Pathologic Findings in Patients With Breast Cancer With BRCA1 and BRCA2 Mutations. *AJR Am J Roentgenol*. 2017;209(4):920-928. doi:10.2214/AJR.16.16957

48. Laenkholm A-V, Soenderstrup IMH, Jensen M-B, et al. Genomic profiling of tumors from patients with germline *BRCA* mutations. *JCO*. 2018;36(15_suppl):1533-1533. doi:10.1200/JCO.2018.36.15_suppl.1533

49. Hansford S, Kaurah P, Li-Chang H, et al. Hereditary Diffuse Gastric Cancer Syndrome: CDH1 Mutations and Beyond. *JAMA oncology*. 2015;1:23-32. doi:10.1001/jamaoncol.2014.168

50. Guilford P, Hopkins J, Harraway J, et al. E-cadherin germline mutations in familial gastric cancer. *Nature*. 1998;392:402-405. doi:10.1038/32918

51. Kaurah P, Huntsman DG. Hereditary Diffuse Gastric Cancer. In: Adam MP, Ardinger HH, Pagon RA, et al., eds. *GeneReviews®*. Seattle (WA): University of Washington, Seattle; 2019. http://www.ncbi.nlm.nih.gov/books/NBK1139/. Accessed June 10, 2019.

52. Zhang H, Feng M, Feng Y, et al. Germline mutations in hereditary diffuse gastric cancer. *Chinese Journal of Cancer Research*. 2018;30(1):122-130. doi:10.21147/j.issn.1000-9604.2018.01.13

53. van der Post RS, Vogelaar IP, Carneiro F, et al. Hereditary diffuse gastric cancer: updated clinical guidelines with an emphasis on germline CDH1 mutation carriers. *Journal of medical genetics*. 2015;52:361-374. doi:10.1136/jmedgenet-2015-103094

54. Petridis C, Shinomiya I, Kohut K, et al. Germline CDH1 mutations in bilateral lobular carcinoma in situ. *British journal of cancer*. 2014;110:1053-1057. doi:10.1038/bjc.2013.792

55. Benusiglio PR, Malka D, Rouleau E, et al. CDH1 germline mutations and the hereditary diffuse gastric and lobular breast cancer syndrome: a multicentre study. *Journal of medical genetics*. 2013;50:486-489. doi:10.1136/jmedgenet-2012-101472

56. Corso G, Intra M, Trentin C, Veronesi P, Galimberti V. CDH1 germline mutations and hereditary lobular breast cancer. *Familial cancer*. 2016;15:215-219. doi:10.1007/s10689-016-9869-5

57. Tedaldi G, Tebaldi M, Zampiga V, et al. Multiple-gene panel analysis in a case series of 255 women with hereditary breast and ovarian cancer. *Oncotarget*. April 2017. doi:10.18632/oncotarget.16791

58. Kraus C, Hoyer J, Vasileiou G, et al. Gene panel sequencing in familial breast/ovarian cancer patients identifies multiple novel mutations also in genes others than BRCA1/2. *Int J Cancer*. 2017;140(1):95-102. doi:10.1002/ijc.30428

59. Kleibl Z, Kristensen VN. Women at high risk of breast cancer: Molecular characteristics, clinical presentation and management. *Breast*. 2016;28:136-144. doi:10.1016/j.breast.2016.05.006

60. Shimelis H, LaDuca H, Hu C, et al. Triple-Negative Breast Cancer Risk Genes Identified by Multigene Hereditary Cancer Panel Testing. *J Natl Cancer Inst*. August 2018. doi:10.1093/jnci/djy106

61. Hahnen E, Hauke J, Engel C, Neidhardt G, Rhiem K, Schmutzler RK. Germline Mutations in Triple-Negative Breast Cancer. *Breast Care (Basel)*. 2017;12(1):15-19. doi:10.1159/000455999

62. Kurian AW, Hughes E, Handorf EA, et al. Breast and Ovarian Cancer Penetrance Estimates Derived From Germline Multiple-Gene Sequencing Results in Women. *JCO Precision Oncology*. 2017;(1):1-12. doi:10.1200/PO.16.00066

63. Antoniou AC, Casadei S, Heikkinen T, et al. Breast-cancer risk in families with mutations in PALB2. *The New England journal of medicine*. 2014;371:497-506. doi:10.1056/NEJMoa1400382

64. Casadei S, Norquist BM, Walsh T, et al. Contribution of inherited mutations in the BRCA2-interacting protein PALB2 to familial breast cancer. *Cancer research*. 2011;71:2222-2229. doi:10.1158/0008-5472.CAN-10-3958

65. Tischkowitz M, Capanu M, Sabbaghian N, et al. Rare germline mutations in PALB2 and breast cancer risk: a population-based study. *Human mutation*. 2012;33:674-680. doi:10.1002/humu.22022

66. Starink TM, van der Veen JP, Arwert F, et al. The Cowden syndrome: a clinical and genetic study in 21 patients. *Clinical genetics*. 1986;29:222-233.

67. Eng C. Will the real Cowden syndrome please stand up: revised diagnostic criteria. *Journal of medical genetics*. 2000;37:828-830.

68. Yehia L, Eng C. 65 YEARS OF THE DOUBLE HELIX: One gene, many endocrine and metabolic syndromes: PTEN-opathies and precision medicine. *Endocr Relat Cancer*. 2018;25(8):T121-T140. doi:10.1530/ERC-18-0162

69. Gammon A, Jasperson K, Champine M. Genetic basis of Cowden syndrome and its implications for clinical practice and risk management. *Appl Clin Genet*. 2016;9:83-92. doi:10.2147/TACG.S41947

70. Ngeow J, Sesock K, Eng C. Breast cancer risk and clinical implications for germline PTEN mutation carriers. *Breast Cancer Res Treat*. 2017;165(1):1-8. doi:10.1007/s10549-015-3665-z

71. Jeghers H, McKUSICK VA, Katz KH. Generalized intestinal polyposis and melanin spots of the oral mucosa, lips and digits; a syndrome of diagnostic significance. *N Engl J Med*. 1949;241(25):993, illust; passim. doi:10.1056/NEJM194912222412501

72. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD. Online Mendelian Inheritance in Man, OMIM. Online Mendelian Inheritance in Man, OMIM. https://omim.org/.

73. van Lier MG, Wagner A, Mathus-Vliegen EM, Kuipers EJ, Steyerberg EW, van Leerdam ME. High cancer risk in Peutz-Jeghers syndrome: a systematic review and surveillance recommendations. *The American journal of gastroenterology*. 2010;105:1258-1264; author reply 1265. doi:10.1038/ajg.2009.725

74. Lim W, Olschwang S, Keller JJ, et al. Relative frequency and morphology of cancers in STK11 mutation carriers. *Gastroenterology*. 2004;126:1788-1794.

75. Malkin D. Li-fraumeni syndrome. *Genes Cancer*. 2011;2(4):475-484. doi:10.1177/1947601911413466

76. Chompret A, Brugieres L, Ronsin M, et al. P53 germline mutations in childhood cancers and cancer risk for carrier individuals. *British journal of cancer*. 2000;82:1932-1937. doi:10.1054/bjoc.2000.1167

77. Nichols KE, Malkin D, Garber JE, Fraumeni JF, Li FP. Germ-line p53 mutations predispose to a wide spectrum of early-onset cancers. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2001;10:83-87.

78. Birch JM, Hartley AL, Tricker KJ, et al. Prevalence and diversity of constitutional mutations in the p53 gene among 21 Li-Fraumeni families. *Cancer Res*. 1994;54(5):1298-1304.

79. Leroy B, Fournier JL, Ishioka C, et al. The TP53 website: an integrative resource centre for the TP53 mutation database and TP53 mutant analysis. *Nucleic acids research*. 2013;41:D962-9. doi:10.1093/nar/gks1033

80. Mai PL, Best AF, Peters JA, et al. Risks of first and subsequent cancers among *TP53* mutation carriers in the National Cancer Institute Li-Fraumeni syndrome cohort: Cancer Risk in *TP53* Mutation Carriers. *Cancer*. 2016;122(23):3673-3681. doi:10.1002/cncr.30248

81. Garber JE, Offit K. Hereditary cancer predisposition syndromes. *J Clin Oncol*. 2005;23:276-292. doi:10.1200/JCO.2005.10.042

82. Walsh T, Casadei S, Coats KH, et al. Spectrum of mutations in BRCA1, BRCA2, CHEK2, and TP53 in families at high risk of breast cancer. *Jama*. 2006;295:1379-1388. doi:10.1001/jama.295.12.1379

83. De Silva S, Tennekoon KH, Karunanayake EH. Overview of the genetic basis toward early detection of breast cancer. *Breast Cancer (Dove Med Press)*. 2019;11:71-80. doi:10.2147/BCTT.S185870

84. Li J, Meeks H, Feng B-J, et al. Targeted massively parallel sequencing of a panel of putative breast cancer susceptibility genes in a large cohort of multiple-case breast and ovarian cancer families. *Journal of Medical Genetics*. 2016;53(1):34-42. doi:10.1136/jmedgenet-2015-103452

85. Goldgar DE, Healey S, Dowty JG, et al. Rare variants in the ATM gene and risk of breast cancer. *Breast cancer research : BCR*. 2011;13:R73. doi:10.1186/bcr2919

86. Couch FJ, Shimelis H, Hu C, et al. Associations between cancer predisposition testing panel genes and breast cancer. *JAMA oncology*. 2017. doi:10.1001/jamaoncol.2017.0424

87. Seal S, Thompson D, Renwick A, et al. Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. *Nature genetics*. 2006;38:1239-1241. doi:10.1038/ng1902

88. Couch FJ, Hart SN, Sharma P, et al. Inherited mutations in 17 breast cancer susceptibility genes among a large triple-negative breast cancer cohort unselected for family history of breast cancer. *J Clin Oncol*. 2015;33:304-311. doi:10.1200/JCO.2014.57.1414

89. Buys SS, Sandbach JF, Gammon A, et al. A study of over 35,000 women with breast cancer tested with a 25-gene panel of hereditary cancer genes. *Cancer*. January 2017. doi:10.1002/cncr.30498

90. Rafnar T, Gudbjartsson DF, Sulem P, et al. Mutations in BRIP1 confer high risk of ovarian cancer. *Nature genetics*. 2011;43:1104-1107. doi:10.1038/ng.955

91. Ramus SJ, Song H, Dicks E, et al. Germline Mutations in the BRIP1, BARD1, PALB2, and NBN Genes in Women With Ovarian Cancer. *Journal of the National Cancer Institute*. 2015;107. doi:10.1093/jnci/djv214

92. Weber-Lassalle N, Hauke J, Ramser J, et al. BRIP1 loss-of-function mutations confer high risk for familial ovarian cancer, but not familial breast cancer. *Breast Cancer Research*. 2018;20(1). doi:10.1186/s13058-018-0935-9

93. Ahn J, Urist M, Prives C. The Chk2 protein kinase. *DNA Repair (Amst)*. 2004;3(8-9):1039-1047. doi:10.1016/j.dnarep.2004.03.033

94. Bell DW, Varley JM, Szydlo TE, et al. Heterozygous germ line hCHK2 mutations in Li-Fraumeni syndrome. *Science*. 1999;286:2528-2531.

95. Weischer M, Bojesen SE, Ellervik C, Tybjærg-Hansen A, Nordestgaard BG. *CHEK2* *1100delC Genotyping for Clinical Assessment of Breast Cancer Risk: Meta-Analyses of 26,000 Patient Cases and 27,000 Controls. *Journal of Clinical Oncology*. 2008;26(4):542-548. doi:10.1200/JCO.2007.12.5922

96. Näslund-Koch C, Nordestgaard BG, Bojesen SE. Increased Risk for Other Cancers in Addition to Breast Cancer for CHEK2*1100delC Heterozygotes Estimated From the Copenhagen General Population Study. *J Clin Oncol*. 2016;34(11):1208-1216. doi:10.1200/JCO.2015.63.3594

97. CHEK2 Breast Cancer Case-Control Consortium. CHEK2*1100delC and susceptibility to breast cancer: a collaborative analysis involving 10,860 breast cancer cases and 9,065 controls from 10 studies. *Am J Hum Genet*. 2004;74(6):1175-1182. doi:10.1086/421251

98. Cybulski C, Gorski B, Huzarski T, et al. CHEK2 is a multiorgan cancer susceptibility gene. *American journal of human genetics*. 2004;75:1131-1135. doi:10.1086/426403

99. Cybulski C, Wokołorczyk D, Jakubowska A, et al. Risk of Breast Cancer in Women With a *CHEK2* Mutation With and Without a Family History of Breast Cancer. *JCO*. 2011;29(28):3747-3752. doi:10.1200/JCO.2010.34.0778

100. Adank MA, Verhoef S, Oldenburg RA, et al. Excess breast cancer risk in first degree relatives of CHEK2 *1100delC positive familial breast cancer cases. *European journal of cancer*. 2013;49:1993-1999. doi:10.1016/j.ejca.2013.01.009

101. Pelttari LM, Khan S, Vuorela M, et al. RAD51B in Familial Breast Cancer. *PLoS ONE*. 2016;11(5):e0153788. doi:10.1371/journal.pone.0153788

102. Song H, Dicks E, Ramus SJ, et al. Contribution of Germline Mutations in the *RAD51B* , *RAD51C* , and *RAD51D* Genes to Ovarian Cancer in the Population. *Journal of Clinical Oncology*. 2015;33(26):2901-2907. doi:10.1200/JCO.2015.61.2408

103. Cerretini R, Mercado G, Morganstein J, et al. Germline pathogenic variants in BRCA1, BRCA2, PALB2 and RAD51C in breast cancer women from Argentina. *Breast Cancer Res Treat*. August 2019. doi:10.1007/s10549-019-05411-9

104. Osher DJ, De Leeneer K, Michils G, et al. Mutation analysis of RAD51D in non-BRCA1/2 ovarian and breast cancer families. *Br J Cancer*. 2012;106(8):1460-1463. doi:10.1038/bjc.2012.87

105. Michailidou K, Hall P, Gonzalez-Neira A, et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nature genetics*. 2013;45:353-361, 361e1-2. doi:10.1038/ng.2563

106. Thomas G, Jacobs KB, Kraft P, et al. A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1). *Nature genetics*. 2009;41:579-584. doi:10.1038/ng.353

107. Orr N, Lemnrau A, Cooke R, et al. Genome-wide association study identifies a common variant in RAD51B associated with male breast cancer risk. *Nature genetics*. 2012;44:1182-1184. doi:10.1038/ng.2417

108. Golmard L, Caux-Moncoutier V, Davy G, et al. Germline mutation in the RAD51B gene confers predisposition to breast cancer. *BMC Cancer*. 2013;13(1). doi:10.1186/1471-2407-13-484

109. Li N, Rowley SM, Thompson ER, et al. Evaluating the breast cancer predisposition role of rare variants in genes associated with low-penetrance breast cancer risk SNPs. *Breast Cancer Research*. 2018;20(1). doi:10.1186/s13058-017-0929-z

110. Walsh T, King M-C. Ten genes for inherited breast cancer. *Cancer Cell*. 2007;11(2):103-105. doi:10.1016/j.ccr.2007.01.010

111. Pennington KP, Swisher EM. Hereditary ovarian cancer: beyond the usual suspects. *Gynecologic oncology*. 2012;124:347-353. doi:10.1016/j.ygyno.2011.12.415

112. Pelttari LM, Heikkinen T, Thompson D, et al. RAD51C is a susceptibility gene for ovarian cancer. *Human molecular genetics*. 2011;20:3278-3288. doi:10.1093/hmg/ddr229

113. LaDuca H, Polley EC, Yussuf A, et al. A clinical guide to hereditary cancer panel testing: evaluation of gene-specific cancer associations and sensitivity of genetic testing criteria in a cohort of 165,000 high-risk patients. *Genet Med*. August 2019. doi:10.1038/s41436-019-0633-8

114. Golmard L, Castéra L, Krieger S, et al. Contribution of germline deleterious variants in the RAD51 paralogs to breast and ovarian cancers. *Eur J Hum Genet*. 2017;25(12):1345-1353. doi:10.1038/s41431-017-0021-2

115. Hauke J, Horvath J, Groß E, et al. Gene panel testing of 5589 BRCA1/2-negative index patients with breast cancer in a routine diagnostic setting: results of the German Consortium for Hereditary Breast and Ovarian Cancer. *Cancer Med*. 2018;7(4):1349-1358. doi:10.1002/cam4.1376

116. Konstanta I, Fostira F, Apostolou P, et al. Contribution of RAD51D germline mutations in breast and ovarian cancer in Greece. *J Hum Genet*. 2018;63(11):1149-1158. doi:10.1038/s10038-018-0498-8

117. Chen X, Li Y, Ouyang T, et al. Associations between RAD51D germline mutations and breast cancer risk and survival in BRCA1/2-negative breast cancers. *Ann Oncol*. 2018;29(10):2046-2051. doi:10.1093/annonc/mdy338

118. Loveday C, Turnbull C, Ruark E, et al. Germline RAD51C mutations confer susceptibility to ovarian cancer. *Nat Genet*. 2012;44(5):475-476; author reply 476. doi:10.1038/ng.2224

119. Easton DF, Pooley KA, Dunning AM, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature*. 2007;447:1087-1093. doi:10.1038/nature05887

120. Garcia-Closas M, Hall P, Nevanlinna H, et al. Heterogeneity of breast cancer associations with five susceptibility loci by clinical and pathological characteristics. *PLoS Genet*. 2008;4(4):e1000054. doi:10.1371/journal.pgen.1000054

121. Broeks A, Schmidt MK, Sherman ME, et al. Low penetrance breast cancer susceptibility loci are associated with specific breast tumor subtypes: findings from the Breast Cancer Association Consortium. *Hum Mol Genet*. 2011;20(16):3289-3303. doi:10.1093/hmg/ddr228

122. Goode EL, Chenevix-Trench G, Song H, et al. A genome-wide association study identifies susceptibility loci for ovarian cancer at 2q31 and 8q24. *Nat Genet*. 2010;42(10):874-879. doi:10.1038/ng.668

123. Milne RL, Gaudet MM, Spurdle AB, et al. Assessing interactions between the associations of common genetic susceptibility variants, reproductive history and body mass index with breast cancer risk in the breast cancer association consortium: a combined case-control study. *Breast cancer research : BCR*. 2010;12:R110. doi:10.1186/bcr2797

124. Shuen AY, Foulkes WD. Inherited mutations in breast cancer genes--risk and response. *J Mammary Gland Biol Neoplasia*. 2011;16(1):3-15. doi:10.1007/s10911-011-9213-5

125. Antoniou AC, Easton DF. Polygenic inheritance of breast cancer: Implications for design of association studies. *Genet Epidemiol*. 2003;25(3):190-202. doi:10.1002/gepi.10261

126. Antoniou AC, Pharoah PDP, McMullan G, et al. A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. *British Journal of Cancer*. 2002;86:76-83. doi:10.1038/sj/bjc/6600008

127. Welsh JL, Hoskin TL, Day CN, et al. Clinical Decision-Making in Patients with Variant of Uncertain Significance in BRCA1 or BRCA2 Genes. *Ann Surg Oncol*. 2017;24(10):3067-3072. doi:10.1245/s10434-017-5959-3

128. Nielsen FC, van Overeem Hansen T, Sorensen CS. Hereditary breast and ovarian cancer: new genes in confined pathways. *Nature reviews Cancer*. 2016;16:599-612. doi:10.1038/nrc.2016.72

129. Gracia-Aznarez FJ, Fernandez V, Pita G, et al. Whole Exome Sequencing Suggests Much of Non-BRCA1/BRCA2 Familial Breast Cancer Is Due to Moderate and Low Penetrance Susceptibility Alleles. *Plos One*. 2013;8. doi:10.1371/journal.pone.0055681

130. Lupski JR, Gonzaga-Jauregui C, Yang Y, et al. Exome sequencing resolves apparent incidental findings and reveals further complexity of SH3TC2 variant alleles causing Charcot-Marie-Tooth neuropathy. *Genome medicine*. 2013;5:57. doi:gm461 [pii] 10.1186/gm461

131. Susswein LR, Marshall ML, Nusbaum R, et al. Pathogenic and likely pathogenic variant prevalence among the first 10,000 patients referred for next-generation cancer panel testing. *Genetics in medicine : official journal of the American College of Medical Genetics*. 2016;18:823-832. doi:10.1038/gim.2015.166

132. LaDuca H, Stuenkel AJ, Dolinsky JS, et al. Utilization of multigene panels in hereditary cancer predisposition testing: analysis of more than 2,000 patients. *Genetics in medicine : official journal of the American College of Medical Genetics*. 2014;16:830-837. doi:10.1038/gim.2014.40

133. Snape K, Ruark E, Tarpey P, et al. Predisposition gene identification in common cancers by exome sequencing: insights from familial breast cancer. *Breast Cancer Research and Treatment*. 2012;134:429-433. doi:10.1007/s10549-012-2057-x

134. Jones S, Hruban RH, Kamiyama M, et al. Exomic sequencing identifies PALB2 as a pancreatic cancer susceptibility gene. *Science*. 2009;324:217. doi:10.1126/science.1171202

135. Comino-Mendez I, Gracia-Aznarez FJ, Schiavi F, et al. Exome sequencing identifies MAX mutations as a cause of hereditary pheochromocytoma. *Nature genetics*. 2011;43:663-667. doi:10.1038/ng.861

136. Palles C, Cazier JB, Howarth KM, et al. Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. *Nature genetics*. 2013;45:136-144. doi:10.1038/ng.2503

137. Pan Z-W, Wang X-J, Chen T, et al. Deleterious Mutations in DNA Repair Gene FANCC Exist in BRCA1/2-Negative Chinese Familial Breast and/or Ovarian Cancer Patients. *Front Oncol*. 2019;9:169. doi:10.3389/fonc.2019.00169

138. Thompson ER, Doyle MA, Ryland GL, et al. Exome sequencing identifies rare deleterious mutations in DNA repair genes FANCC and BLM as potential breast cancer susceptibility alleles. *PLoS genetics*. 2012;8:e1002894. doi:10.1371/journal.pgen.1002894

139. Kiiski JI, Pelttari LM, Khan S, et al. Exome sequencing identifies FANCM as a susceptibility gene for triple-negative breast cancer. *Proceedings of the National Academy of Sciences of the United States of America*. 2014;111:15172-15177. doi:10.1073/pnas.1407909111

140. Estivill X, Armengol L. Copy number variants and common disorders: filling the gaps and exploring complexity in genome-wide association studies. *PLoS genetics*. 2007;3:1787-1799. doi:10.1371/journal.pgen.0030190

141. Ruano Y, Mollejo M, de Lope AR, Hernandez-Moneo JL, Martinez P, Melendez B. Microarray-based comparative genomic hybridization (array-CGH) as a useful tool for identifying genes involved in Glioblastoma (GB). *Methods Mol Biol*. 2010;653:35-45. doi:10.1007/978-1-60761-759-4_3

142. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer genome landscapes. *Science*. 2013;339:1546-1558. doi:10.1126/science.1235122

143. Sabatier R, Adélaïde J, Finetti P, et al. BARD1 homozygous deletion, a possible alternative to BRCA1 mutation in basal breast cancer. *Genes Chromosomes Cancer*. 2010;49(12):1143-1151. doi:10.1002/gcc.20822

144. Byrum AK, Vindigni A, Mosammaparast N. Defining and Modulating 'BRCAness.' *Trends in Cell Biology*. 2019;29(9):740-751. doi:10.1016/j.tcb.2019.06.005

145. Wu J-Y, Vlastos A-T, Pelte M-F, et al. Aberrant expression of BARD1 in breast and ovarian cancers with poor prognosis. *Int J Cancer*. 2006;118(5):1215-1226. doi:10.1002/ijc.21428

146. Tanic M, Andres E, Rodriguez-Pinilla SM, et al. MicroRNA-based molecular classification of non-BRCA1/2 hereditary breast tumours. *British journal of cancer*. 2013;109:2724-2734. doi:10.1038/bjc.2013.612

147. Honrado E, Osorio A, Milne RL, et al. Immunohistochemical classification of non-BRCA1/2 tumors identifies different groups that demonstrate the heterogeneity of BRCAX families. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc*. 2007;20:1298-1306. doi:10.1038/modpathol.3800969

148. Gronwald J, Jauch A, Cybulski C, et al. Comparison of genomic abnormalities between BRCAX and sporadic breast cancers studied by comparative genomic hybridization. *International journal of cancer*. 2005;114:230-236. doi:10.1002/ijc.20723

149. Didraga M, Van Beers E, Brandwijk K, et al. Array CGH aberrations in non-BRCA1/2 familial breast cancer. *Annals of Oncology*. 2010;21:iv29-iv30. doi:10.1093/annonc/mdq141

150. Didraga MA, Van Beers EH, Joosse SA, et al. A non-BRCA1/2 hereditary breast cancer sub-group defined by aCGH profiling of genetically related patients. *Breast cancer research and treatment*. 2011;130:425-436. doi:10.1007/s10549-011-1357-x

151. Alvarez C, Aravena A, Tapia T, et al. Different Array CGH profiles within hereditary breast cancer tumors associated to BRCA1 expression and overall survival. *BMC cancer*. 2016;16. doi:10.1186/s12885-016-2261-x

152. Hilbers FS, Meijers CM, Laros JF, et al. Exome sequencing of germline DNA from non-BRCA1/2 familial breast cancer cases selected on the basis of aCGH tumor profiling. *PloS one*. 2013;8:e55734. doi:10.1371/journal.pone.0055734

153. Curtin NJ. DNA repair dysregulation from cancer driver to therapeutic target. *Nature Reviews Cancer*. 2012;12(12):801-817. doi:10.1038/nrc3399

154. Torgovnick A, Schumacher B. DNA repair mechanisms in cancer development and therapy. *Frontiers in Genetics*. 2015;6. doi:10.3389/fgene.2015.00157

155. Aloraifi F. Detection of novel germline mutations for breast cancer in non-BRCA1/2 families. *FEBS Journal*. 2016;283:5. doi:10.1111/febs.13803

156. Torrezan GT, de Almeida FG dos SR, Figueiredo MCP, et al. Complex Landscape of Germline Variants in Brazilian Patients With Hereditary and Early Onset Breast Cancer. *Frontiers in Genetics*. 2018;9. doi:10.3389/fgene.2018.00161

157. Esteban-Jurado C, Vila-Casadesús M, Garre P, et al. Whole-exome sequencing identifies rare pathogenic variants in new predisposition genes for familial colorectal cancer. *Genet Med*. 2015;17(2):131-142. doi:10.1038/gim.2014.89

158. Ten Broeke SW, van Bavel TC, Jansen AML, et al. Molecular Background of Colorectal Tumors From Patients With Lynch Syndrome Associated With Germline Variants in PMS2. *Gastroenterology*. 2018;155(3):844-851. doi:10.1053/j.gastro.2018.05.020

159. Schneider M, Chandler K, Tischkowitz M, Meyer S. Fanconi anaemia: genetics, molecular biology, and cancer - implications for clinical management in children and adults. *Clinical Genetics*. 2015;88(1):13-24. doi:10.1111/cge.12517

160. Weon JL, Glass DA. Novel therapeutic approaches to xeroderma pigmentosum. *British Journal of Dermatology*. November 2018. doi:10.1111/bjd.17253

161. Neamatzadeh H, Shiryazdi SM, Kalantar SM. BRCA1 and BRCA2 mutations in Iranian breast cancer patients: A systematic review. *J Res Med Sci*. 2015;20(3):284-293.

162. Slavin TP, Maxwell KN, Lilyquist J, et al. Author Correction: The contribution of pathogenic variants in breast cancer susceptibility genes to familial breast cancer risk. *npj Breast Cancer*. 2017;3(1):44. doi:10.1038/s41523-017-0046-2

163. Mavaddat N, Michailidou K, Dennis J, et al. Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. *Am J Hum Genet*. 2019;104(1):21-34. doi:10.1016/j.ajhg.2018.11.002

164. Janatova M, Soukupova J, Stribrna J, et al. Mutation Analysis of the RAD51C and RAD51D Genes in High-Risk Ovarian Cancer Patients and Families from the Czech Republic. Brusgaard K, ed. *PLoS ONE*. 2015;10(6):e0127711. doi:10.1371/journal.pone.0127711

165. Lu H-M, Li S, Black MH, et al. Association of Breast and Ovarian Cancers With Predisposition Genes Identified by Large-Scale Sequencing. *JAMA Oncology*. August 2018. doi:10.1001/jamaoncol.2018.2956

166. Palmero EI, Galvão HCR, Fernandes GC, et al. Oncogenetics service and the Brazilian public health system: the experience of a reference Cancer Hospital. *Genetics and Molecular Biology*. 2016;39(2):168-177. doi:10.1590/1678-4685-gmb-2014-0364

167. Das R, Ghosh SK. Genetic variants of the DNA repair genes from Exome Aggregation Consortium (EXAC) database: significance in cancer. *DNA Repair*. 2017;52:92-102. doi:10.1016/j.dnarep.2017.02.013

168. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536:285.

169. Naslavsky MS, Yamamoto GL, de Almeida TF, et al. Exomic variants of an elderly cohort of Brazilians in the ABraOM database: NASLAVSKY ET AL. *Human Mutation*. 2017;38(7):751-763. doi:10.1002/humu.23220

170. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405-424. doi:10.1038/gim.2015.30

171. Kalia SS, Adelman K, Bale SJ, et al. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of

the American College of Medical Genetics and Genomics. *Genet Med*. 2017;19(2):249-255. doi:10.1038/gim.2016.190

172. Heeke AL, Pishvaian MJ, Lynce F, et al. Prevalence of Homologous Recombination-Related Gene Mutations Across Multiple Cancer Types. *JCO Precis Oncol*. 2018;2018. doi:10.1200/PO.17.00286

173. Romero-Laorden N, Castro E. Inherited mutations in DNA repair genes and cancer risk. *Current Problems in Cancer*. 2017;41(4):251-264. doi:10.1016/j.currproblcancer.2017.02.009

174. AlDubayan SH, Giannakis M, Moore ND, et al. Inherited DNA-Repair Defects in Colorectal Cancer. *The American Journal of Human Genetics*. 2018;102(3):401-414. doi:10.1016/j.ajhg.2018.01.018

175. den Brok WD, Schrader KA, Sun S, et al. Homologous Recombination Deficiency in Breast Cancer: A Clinical Review. *JCO Precision Oncology*. 2017;(1):1-13. doi:10.1200/PO.16.00031

176. Daly MB, Pilarski R, Berry M, et al. NCCN Guidelines Insights: Genetic/Familial High-Risk Assessment: Breast and Ovarian, Version 2.2017. *Journal of the National Comprehensive Cancer Network : JNCCN*. 2017;15:9-20.

177. Lee K, Seifert BA, Shimelis H, et al. Clinical validity assessment of genes frequently tested on hereditary breast and ovarian cancer susceptibility sequencing panels. *Genet Med*. 2019;21(7):1497-1506. doi:10.1038/s41436-018-0361-5

178. Kleiblova P, Stolarova L, Krizova K, et al. Identification of deleterious germline CHEK2 mutations and their association with breast and ovarian cancer. *Int J Cancer*. 2019;145(7):1782-1797. doi:10.1002/ijc.32385

179. Southey MC, Goldgar DE, Winqvist R, et al. *PALB2* , *CHEK2* and *ATM* rare variants and cancer risk: data from COGS. *J Med Genet*. 2016;53(12):800-811. doi:10.1136/jmedgenet-2016-103839

180. Matsuda M, Miyagawa K, Takahashi M, et al. Mutations in the RAD54 recombination gene in primary cancers. *Oncogene*. 1999;18(22):3427-3430. doi:10.1038/sj.onc.1202692

181. Shagimardanova E, Shigapova LH, Enikeev RF, et al. 240PLandscape of germline mutations in hereditary breast and ovarian cancer (HBOC) patients in Russia revealed by target panel sequencing. *Annals of Oncology*. 2018;29(suppl_8). doi:10.1093/annonc/mdy270.235

182. Roberts ME, Jackson SA, Susswein LR, et al. MSH6 and PMS2 germ-line pathogenic variants implicated in Lynch syndrome are associated with breast cancer. *Genetics in Medicine*. 2018;20(10):1167-1174. doi:10.1038/gim.2017.254

183. Smith AL, Alirezaie N, Connor A, et al. Candidate DNA repair susceptibility genes identified by exome sequencing in high-risk pancreatic cancer. *Cancer Lett*. 2016;370(2):302-312. doi:10.1016/j.canlet.2015.10.030

184. Lachaud C, Moreno A, Marchesi F, Toth R, Blow JJ, Rouse J. Ubiquitinated Fancd2 recruits Fan1 to stalled replication forks to prevent genome instability. *Science*. 2016;351(6275):846-849. doi:10.1126/science.aad5634

185. Vargas G, Dámaso E, Navarro M, et al. Identification of germline FAN1 variants in MSH2-deficient Lynch-like syndrome patients. April 2016. https://ehtg.org/2016/04/p19-identification-germline-fan1-variants-msh2-deficient-lynch-like-syndrome-patients/.

186. Seguí N, Mina LB, Lázaro C, et al. Germline Mutations in FAN1 Cause Hereditary Colorectal Cancer by Impairing DNA Repair. *Gastroenterology*. 2015;149(3):563-566. doi:10.1053/j.gastro.2015.05.056

187. Park DJ, Odefrey FA, Hammet F, et al. FAN1 variants identified in multiple-case early-onset breast cancer families via exome sequencing: no evidence for association with risk for breast cancer. *Breast cancer research and treatment*. 2011;130:1043-1049. doi:10.1007/s10549-011-1704-y

188. Kanchi KL, Johnson KJ, Lu C, et al. Integrated analysis of germline and somatic variants in ovarian cancer. *Nat Commun*. 2014;5:3156. doi:10.1038/ncomms4156

189. Walsh CS. Two decades beyond BRCA1/2: Homologous recombination, hereditary cancer risk and a target for ovarian cancer therapy. *Gynecologic Oncology*. 2015;137(2):343-350. doi:10.1016/j.ygyno.2015.02.017

190. Mylavarapu S, Das A, Roy M. Role of BRCA Mutations in the Modulation of Response to Platinum Therapy. *Front Oncol*. 2018;8:16. doi:10.3389/fonc.2018.00016

191. Stafford JL, Dyson G, Levin NK, et al. Reanalysis of BRCA1/2 negative high risk ovarian cancer patients reveals novel germline risk loci and insights into missing heritability. *Plos One*. 2017;12. doi:10.1371/journal.pone.0178450

192. Ceccaldi R, Liu JC, Amunugama R, et al. Homologous-recombination-deficient tumours are dependent on Polθ-mediated repair. *Nature*. 2015;518(7538):258-262. doi:10.1038/nature14184

193. Andressoo J-O, Hoeijmakers JHJ, Mitchell JR. Nucleotide excision repair disorders and the balance between cancer and aging. *Cell Cycle*. 2006;5(24):2886-2888. doi:10.4161/cc.5.24.3565

194. Lynch HT, Lynch J, Conway T, et al. Hereditary breast cancer and family cancer syndromes. *World J Surg*. 1994;18(1):21-31. doi:10.1007/bf00348188

195. Bodily WR, Shirts BH, Walsh T, et al. *Markers of BRCAness in Breast Cancer*. Cancer Biology; 2019. doi:10.1101/662874

196. Akashi-Tanaka S, Watanabe C, Takamaru T, et al. BRCAness Predicts Resistance to Taxane-Containing Regimens in Triple Negative Breast Cancer During Neoadjuvant Chemotherapy. *Clinical Breast Cancer*. 2015;15(1):80-85. doi:10.1016/j.clbc.2014.08.003

197. Tanino H, Kosaka Y, Nishimiya H, et al. BRCAness and Prognosis in Triple-Negative Breast Cancer Patients Treated with Neoadjuvant Chemotherapy. *PLoS ONE*. 2016;11(12):e0165721. doi:10.1371/journal.pone.0165721

198. Turner N, Tutt A, Ashworth A. Hallmarks of “BRCAness” in sporadic cancers. *Nat Rev Cancer*. 2004;4(10):814-819. doi:10.1038/nrc1457

199. Mori H, Kubo M, Kai M, et al. BRCAness Combined With a Family History of Cancer Is Associated With a Poor Prognosis for Breast Cancer Patients With a High Risk of BRCA Mutations. *Clinical Breast Cancer*. 2018;18(5):e1217-e1227. doi:10.1016/j.clbc.2018.05.008

200. Sharma P. Biology and Management of Patients With Triple-Negative Breast Cancer. *Oncologist*. 2016;21(9):1050-1062. doi:10.1634/theoncologist.2016-0067

201. Rigakos G, Razis E. BRCAness: finding the Achilles heel in ovarian cancer. *Oncologist*. 2012;17(7):956-962. doi:10.1634/theoncologist.2012-0028

202. Hjortkjær M, Malik Aagaard Jørgensen M, Waldstrøm M, et al. The clinical importance of BRCAness in a population-based cohort of Danish epithelial ovarian cancer. *Int J Gynecol Cancer*. 2019;29(1):166-173. doi:10.1136/ijgc-2018-000017

203. Kerber RA, O'Brien E. A cohort study of cancer risk in relation to family histories of cancer in the Utah population database. *Cancer*. 2005;103(9):1906-1915. doi:10.1002/cncr.20989

204. Poole CA, Byers T, Calle EE, Bondy J, Fain P, Rodriguez C. Influence of a family history of cancer within and across multiple sites on patterns of cancer mortality risk for women. *Am J Epidemiol*. 1999;149(5):454-462. doi:10.1093/oxfordjournals.aje.a009833

205. Slattery ML, Kerber RA. A comprehensive evaluation of family history and breast cancer risk. The Utah Population Database. *JAMA*. 1993;270(13):1563-1568.

206. Hemminki K, Sundquist J, Lorenzo Bermejo J. Familial risks for cancer as the basis for evidence-based clinical referral and counseling. *Oncologist*. 2008;13(3):239-247. doi:10.1634/theoncologist.2007-0242

207. Lee J-Y, Kim J, Kim S-W, et al. BRCA1/2-negative, high-risk breast cancers (BRCAX) for Asian women: genetic susceptibility loci and their potential impacts. *Sci Rep*. 2018;8(1):15263. doi:10.1038/s41598-018-31859-8

208. Mangia A, Chiarappa P, Tommasi S, et al. Genetic heterogeneity by comparative genomic hybridization in BRCAx breast cancers. *Cancer Genetics and Cytogenetics*. 2008;182:75-83. doi:10.1016/j.cancergencyto.2008.01.002

209. Letessier A, Sircoulomb F, Ginestier C, et al. Frequency, prognostic impact, and subtype association of 8p12, 8q24, 11q13, 12p13, 17q12, and 20q13 amplifications in breast cancers. *BMC Cancer*. 2006;6(1). doi:10.1186/1471-2407-6-245

210. Yang ZQ, Streicher KL, Ray ME, Abrams J, Ethier SP. Multiple Interacting Oncogenes on the 8p11-p12 Amplicon in Human Breast Cancer. *Cancer Research*. 2006;66(24):11632-11643. doi:10.1158/0008-5472.CAN-06-2946

211. Bernard-Pierrot I, Gruel N, Stransky N, et al. Characterization of the recurrent 8p11-12 amplicon identifies PPAPDC1B, a phosphatase protein, as a new therapeutic target in breast cancer. *Cancer Res*. 2008;68(17):7165-7175. doi:10.1158/0008-5472.CAN-08-1360

212. Gelsi-Boyer V. Comprehensive Profiling of 8p11-12 Amplification in Breast Cancer. *Molecular Cancer Research*. 2005;3(12):655-667. doi:10.1158/1541-7786.MCR-05-0128

213. Rhodes DR, Yu J, Shanker K, et al. ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia*. 2004;6(1):1-6. doi:10.1016/s1476-5586(04)80047-2

214. Shao F, Sun H, Deng C-X. Potential therapeutic targets of triple-negative breast cancer based on its intrinsic subtype. *Oncotarget*. 2017;8(42):73329-73344. doi:10.18632/oncotarget.20274

215. Porta R, Borea R, Coelho A, et al. FGFR a promising druggable target in cancer: Molecular biology and new drugs. *Crit Rev Oncol Hematol*. 2017;113:256-267. doi:10.1016/j.critrevonc.2017.02.018

216. Chen H, Singh RR, Lu X, et al. Genome-wide copy number aberrations and HER2 and FGFR1 alterations in primary breast cancer by molecular inversion probe microarray. *Oncotarget*. 2017;8(7):10845-10857. doi:10.18632/oncotarget.14802

217. Luo J, Liu S, Leung S, et al. An mRNA Gene Expression-Based Signature to Identify FGFR1-Amplified Estrogen Receptor-Positive Breast Tumors. *J Mol Diagn*. 2017;19(1):147-161. doi:10.1016/j.jmoldx.2016.09.007

218. Shi Y-J, Tsang JYS, Ni Y-B, Chan S-K, Chan K-F, Tse GM. FGFR1 is an adverse outcome indicator for luminal A breast cancers. *Oncotarget*. 2016;7(4):5063-5073. doi:10.18632/oncotarget.6563

219. André F, Bachelot T, Campone M, et al. Targeting FGFR with dovitinib (TKI258): preclinical and clinical data in breast cancer. *Clin Cancer Res*. 2013;19(13):3693-3702. doi:10.1158/1078-0432.CCR-13-0190

220. ClinicalTrials. https://clinicaltrials.gov/.

221. Lundgren K, Holm K, Nordenskjöld B, Borg Å, Landberg G. Gene products of chromosome 11q and their association with CCND1gene amplification and tamoxifen resistance in premenopausal breast cancer. *Breast Cancer Res*. 2008;10(5):R81. doi:10.1186/bcr2150

222. Sheen-Chen S-M, Huang C-Y, Liu Y-Y, Huang C-C, Tang R-P. Cortactin in breast cancer: analysis with tissue microarray. *Anticancer Res*. 2011;31(1):293-297.

223. Tasian SK, Loh ML. Understanding the biology of CRLF2-overexpressing acute lymphoblastic leukemia. *Crit Rev Oncog*. 2011;16(1-2):13-24.

224. Jacot W, Fiche M, Zaman K, Wolfer A, Lamy P-J. The HER2 amplicon in breast cancer: Topoisomerase IIA and beyond. *Biochim Biophys Acta*. 2013;1836(1):146-157. doi:10.1016/j.bbcan.2013.04.004

225. Cliby W, Ritland S, Hartmann L, et al. Human epithelial ovarian cancer allelotype. *Cancer Res*. 1993;53(10 Suppl):2393-2398.

226. Castells A, Ino Y, Louis DN, Ramesh V, Gusella JF, Rustgi AK. Mapping of a target region of allelic loss to a 0.5-cM interval on chromosome 22q13 in human colorectal cancer. *Gastroenterology*. 1999;117(4):831-837.

227. Zweemer RP, Ryan A, Snijders AM, et al. Comparative genomic hybridization of microdissected familial ovarian carcinoma: two deleted regions on chromosome 15q not previously identified in sporadic ovarian carcinoma. *Lab Invest*. 2001;81(10):1363-1370.

228. Chong GO, Jeon H-S, Han HS, et al. Differential MicroRNA Expression Profiles in Primary and Recurrent Epithelial Ovarian Cancer. *Anticancer Res*. 2015;35(5):2611-2617.

229. Alemar B, Herzog J, Brinckmann Oliveira Netto C, et al. Prevalence of Hispanic BRCA1 and BRCA2 mutations among hereditary breast and ovarian cancer patients from Brazil reveals differences among Latin American populations. *Cancer Genet*. 2016;209(9):417-422. doi:10.1016/j.cancergen.2016.06.008

230. Alemar B, Gregório C, Herzog J, et al. BRCA1 and BRCA2 mutational profile and prevalence in hereditary breast and ovarian cancer (HBOC) probands from Southern Brazil: Are international testing criteria appropriate for this specific population? *PLoS ONE*. 2017;12(11):e0187630. doi:10.1371/journal.pone.0187630

231. Kast K, Rhiem K, Wappenschmidt B, et al. Prevalence of BRCA1/2 germline mutations in 21 401 families with breast and ovarian cancer. *J Med Genet*. 2016;53(7):465-471. doi:10.1136/jmedgenet-2015-103672

232. Ruiz-Linares A, Adhikari K, Acuña-Alonzo V, et al. Admixture in Latin America: Geographic Structure, Phenotypic Diversity and Self-Perception of Ancestry Based on 7,342 Individuals. Di Rienzo A, ed. *PLoS Genet*. 2014;10(9):e1004572. doi:10.1371/journal.pgen.1004572

233. Kehdy FSG, Gouveia MH, Machado M, et al. Origin and dynamics of admixture in Brazilians and its effect on the pattern of deleterious mutations. *Proc Natl Acad Sci USA*. 2015;112(28):8696-8701. doi:10.1073/pnas.1504447112

234. Ewald IP, Cossio SL, Palmero EI, et al. BRCA1 and BRCA2 rearrangements in Brazilian individuals with Hereditary Breast and Ovarian Cancer Syndrome. *Genet Mol Biol*. 2016;39(2):223-231. doi:10.1590/1678-4685-GMB-2014-0350

235. Moreira MAM, Bobrovnitchaia IG, Lima MAFD, et al. Portuguese c.156_157insAlu BRCA2 founder mutation: gastrointestinal and tongue neoplasias may be part of the phenotype. *Fam Cancer*. 2012;11(4):657-660. doi:10.1007/s10689-012-9551-5

236. Forbes SA, Beare D, Boutselakis H, et al. COSMIC: somatic cancer genetics at high-resolution. *Nucleic Acids Research*. 2017;45(D1):D777-D783. doi:10.1093/nar/gkw1121

237. Karczewski KJ, Francioli LC, Tiao G, et al. *Variation across 141,456 Human Exomes and Genomes Reveals the Spectrum of Loss-of-Function Intolerance across Human Protein-Coding Genes*. Genomics; 2019. doi:10.1101/531210

238. GeneCards. https://www.genecards.org/.

239. Qian F, Feng Y, Zheng Y, et al. Genetic variants in microRNA and microRNA biogenesis pathway genes and breast cancer risk among women of African ancestry. *Hum Genet*. 2016;135(10):1145-1159. doi:10.1007/s00439-016-1707-1

240. Torrezan GT, Ferreira EN, Nakahata AM, et al. Recurrent somatic mutation in DROSHA induces microRNA profile changes in Wilms tumour. *Nat Commun*. 2014;5:4039. doi:10.1038/ncomms5039

241. Lynch HT, Snyder CL, Shaw TG, Heinen CD, Hitchins MP. Milestones of Lynch syndrome: 1895–2015. *Nature Reviews Cancer*. 2015;15:181.

242. Buckley AR, Ideker T, Carter H, Harismendy O, Schork NJ. Exome-wide analysis of bi-allelic alterations identifies a Lynch phenotype in The Cancer Genome Atlas. *Genome Med*. 2018;10(1):69. doi:10.1186/s13073-018-0579-5

243. Raskin L, Guo Y, Du L, et al. Targeted sequencing of established and candidate colorectal cancer genes in the Colon Cancer Family Registry Cohort. *Oncotarget*. 2017;8(55):93450-93463. doi:10.18632/oncotarget.18596

244. Wang X, Szabo C, Qian C, et al. Mutational analysis of thirty-two double-strand DNA break repair genes in breast and pancreatic cancers. *Cancer Res*. 2008;68(4):971-975. doi:10.1158/0008-5472.CAN-07-6272

245. Brandalize APC, Schüler-Faccini L, Hoffmann J-S, Caleffi M, Cazaux C, Ashton-Prolla P. A DNA repair variant in POLQ (c.-1060A > G) is associated to hereditary breast cancer patients: a case-control study. *BMC Cancer*. 2014;14:850. doi:10.1186/1471-2407-14-850

246. Family L, Bensen JT, Troester MA, Wu MC, Anders CK, Olshan AF. Single-nucleotide polymorphisms in DNA bypass polymerase genes and association with breast cancer and breast cancer subtypes among African Americans and Whites. *Breast Cancer Res Treat*. 2015;149(1):181-190. doi:10.1007/s10549-014-3203-4

247. Kim A, Jang MH, Lee SJ, Bae YK. Mutations of the Epidermal Growth Factor Receptor Gene in Triple-Negative Breast Cancer. *J Breast Cancer*. 2017;20(2):150-159. doi:10.4048/jbc.2017.20.2.150

248. Mavroeidis L, Metaxa-Mariatou V, Papoudou-Bai A, et al. Comprehensive molecular screening by next generation sequencing reveals a distinctive mutational profile of KIT/PDGFRA genes and novel genomic alterations: results from a 20-year cohort of patients with GIST from north-western Greece. *ESMO Open*. 2018;3(3):e000335. doi:10.1136/esmoopen-2018-000335

249. Xu Z, Huo X, Tang C, et al. Frequent KIT mutations in human gastrointestinal stromal tumors. *Sci Rep*. 2014;4:5907. doi:10.1038/srep05907

250. Mishra R, Hanker AB, Garrett JT. Genomic alterations of ERBB receptors in cancer: clinical implications. *Oncotarget*. 2017;8(69):114371-114392. doi:10.18632/oncotarget.22825

251. Donoghue JF, Kerr LT, Alexander NW, et al. Activation of ERBB4 in Glioblastoma Can Contribute to Increased Tumorigenicity and Influence Therapeutic Response. *Cancers (Basel)*. 2018;10(8). doi:10.3390/cancers10080243

252. Penkert J, Schmidt G, Hofmann W, et al. Breast cancer patients suggestive of Li-Fraumeni syndrome: mutational spectrum, candidate genes, and unexplained heredity. *Breast Cancer Research*. 2018;20(1). doi:10.1186/s13058-018-1011-1

253. El-Gebali S, Mistry J, Bateman A, et al. The Pfam protein families database in 2019. *Nucleic Acids Research*. 2019;47(D1):D427-D432. doi:10.1093/nar/gky995

254. Valletta D, Czech B, Spruss T, et al. Regulation and function of the atypical cadherin FAT1 in hepatocellular carcinoma. *Carcinogenesis*. 2014;35(6):1407-1415. doi:10.1093/carcin/bgu054

255. Katoh M. Function and cancer genomics of FAT family genes (review). *Int J Oncol*. 2012;41(6):1913-1918. doi:10.3892/ijo.2012.1669

256. Furukawa T, Sakamoto H, Takeuchi S, et al. Whole exome sequencing reveals recurrent mutations in BRCA2 and FAT genes in acinar cell carcinomas of the pancreas. *Sci Rep*. 2015;5:8829. doi:10.1038/srep08829

257. Connor AA. Germline & somatic variation in gastrointestinal malignancies. 2017. http://hdl.handle.net/1807/80761.

258. Vogelaar IP, van der Post RS, van Krieken JHJ, et al. Unraveling genetic predisposition to familial or early onset gastric cancer using germline whole-exome sequencing. *Eur J Hum Genet*. 2017;25(11):1246-1252. doi:10.1038/ejhg.2017.138

259. Johannsdottir HK, Johannesdottir G, Agnarsson BA, et al. Deletions on chromosome 4 in sporadic and BRCA mutated tumors and association with pathological variables. *Anticancer Res*. 2004;24(5A):2681-2687.

260. Horn S, Figl A, Rachakonda PS, et al. TERT promoter mutations in familial and sporadic melanoma. *Science*. 2013;339(6122):959-961. doi:10.1126/science.1230062

261. Sarek G, Marzec P, Margalef P, Boulton SJ. Molecular basis of telomere dysfunction in human genetic diseases. *Nat Struct Mol Biol*. 2015;22(11):867-874. doi:10.1038/nsmb.3093

ANEXOS

ANEXO I - Arquivo PDF do artigo publicado na revista científica "*Cancer Genetics*".

ELSEVIER

Cancer Genetics 228–229 (2018) 93–97

Cancer Genetics

SHORT COMMUNICATION

Screening and characterization of *BRCA2* c.156_157insAlu in Brazil: Results from 1380 individuals from the South and Southeast

Paula Silva Felicio[a,1], Barbara Alemar[b,c,1], Aline Silva Coelho[a], Gustavo Noriz Berardinelli[a], Matias Eliseo Melendez[a], André Van Helvoort Lengert[a], Rodrigo Depieri Miche Ili[d], Rui M. Reis[a,e,f], Gabriela Carvalho Fernandes[a], Ingrid Petroni Ewald[g], Camila Matzenbacher Bittar[b,c], Cristina Brinckmann Oliveira Netto[h], Osvaldo Artigalas[i], Ana Peixoto[j], Manuela Pinheiro[j], Manuel R. Teixeira[j,k], Fernando Regla Vargas[l,m], Anna Cláudia Evangelista dos Santos[n], Miguel Angelo Martins Moreira[n], Patricia Ashton-Prolla[b,c,h,2,*], Edenir Inêz Palmero[a,o,2,**]

[a] *Molecular Oncology Research Center, Barretos Cancer Hospital, R. Antenor Duarte Villela, 1331. Barretos 14784-400, São Paulo, Brazil;* [b] *Programa de Pós-graduação em Genética e Biologia Molecular, Universidade Federal do Rio Grande do Sul. Av. Bento Gonçalves, 9500 – Prédio 43323M. Porto Alegre 91501-970, Rio Grande do Sul, Brazil;* [c] *Laboratório de Medicina Genômica—Hospital de Clínicas de Porto Alegre (HCPA), Rua Ramiro Barcelos, 2350. Porto Alegre 90035-903, Rio Grande do Sul, Brazil;* [d] *Department of Oncogenetics, Barretos Cancer Hospital, R. Antenor Duarte Villela, 1331. Barretos 14784-400, São Paulo, Brazil;* [e] *Life and Health Sciences Research Institute (ICVS), School of Health Sciences, University of Minho, Campus de Gualtar, Braga 4710-057, Portugal;* [f] *ICVS/3B's—PT Government Associate Laboratory, Braga/Guimarães, Portugal;* [g] *Santa Casa de Misericórdia de Porto Alegre. Av. Independência, 75. Porto Alegre 90035-072, Rio Grande do Sul, Brazil;* [h] *Serviço de Genética Médica, Hospital de Clínicas de Porto Alegre. Rua Ramiro Barcelos, 2350. Porto Alegre 90035-903, Rio Grande do Sul, Brazil;* [i] *Hospital Moinhos de Vento, Rua Ramiro Barcelos, 910. Porto Alegre 91790-560, Rio Grande do Sul, Brazil;* [j] *Department of Genetics, Portuguese Oncology Institute, R. Dr António Bernardino de Almeida. Porto 4200-072, Portugal;* [k] *Biomedical Sciences Institute Abel Salazar (ICBAS), University of Porto, R. Jorge de Viterbo Ferreira, 228, Porto 4050-313, Portugal;* [l] *Laboratório de Epidemiologia de Malformações Congênitas, Instituto Oswaldo Cruz (FIOCRUZ), Brazil;* [m] *Universidade Federal do Estado do Rio de Janeiro, Brazil;* [n] *Genetics Program, Instituto Nacional de Câncer, Rua André Cavalcanti, 37. Rio de Janeiro 20231-050, Rio de Janeiro, Brazil;* [o] *Barretos School of Health Science, Dr. Paulo Prata. Av. Loja Maçonica Renovadora, 68. Barretos 14785-002, São Paulo, Brazil*

Received March 20, 2018; received in revised form July 31, 2018; accepted September 17, 2018

* Corresponding author. Present address: Serviço de Genética Médica, Hospital de Clínicas de Porto Alegre. Rua Ramiro Barcelos, 2350. Porto Alegre 90035-903, Rio Grande do Sul, Brazil. Email: pprolla@hcpa.edu.br

** Corresponding author at: Centro de Pesquisa em Oncologia Molecular, Hospital de Câncer de Barretos, Avenida Antenor Duarte Vilela, 1331, Barretos 14784-400, São Paulo, Brazil. Email: edenirip@yahoo.com.br

E-mail addresses: feliciops@hotmail.com, barbara.alemar@gmail.com, alinecoelhobiomed@gmail.com, gnoriz@gmail.com, ahlengert@gmail.com, ruireis.hcb@gmail.com, gabriela.carvalhofernandes@gmail.com, ipewald@gmail.com, camilambittar@gmail.com, cbnetto@hcpa.edu.br, artigalas@gmail.com, anpeixoto@gmail.com, manuelap@ipoporto.min-saude.pt, manuelteixeira@ipoporto.min-saude.pt, miguelm@inca.gov.br, pprolla@gmail.com, edenirip@yahoo.com.br

[1] Both authors contributed equally to this work.

[2] Both authors jointly supervised this work.

https://doi.org/10.1016/j.cancergen.2018.09.001

Abstract

Portuguese immigration to Brazil occurred in several waves and greatly contributed to the genetic composition of current Brazilian population. In this study, we evaluated the frequency of a Portuguese founder Alu insertion in *BRCA2* exon 3 (c.156_157insAlu) among individuals fulfilling Hereditary Breast and Ovarian Cancer (HBOC) syndrome criteria in 1,380 unrelated families originated from three distinct Brazilian States. We identified the c.156_157insAlu *BRCA2* mutation in nine (9/1,380; 0.65%) probands analised. In carrier probands, European ancestry had the highest proportion (80%), followed by the African (10%) and Amerindian and in most families with the rearrangement, haplotype analyses were compatible with the Portuguese ancestral haplotype. In conclusion, the present study reports a low albeit relevant frequency of the Portuguese *BRCA2* founder mutation c.156_157insAlu in Brazilian patients at-risk for HBOC Brazilian population.

Keywords BRCA2, HBOC, Genetic screening, Alu elements.

Introduction

Alu elements are retroelements (retrotransposons) typically 300 nt in size, which account for approximately 11% of the human genome. It is believed that genomic rearrangements induced by Alu elements account for approximately 0.1% of human disease [1,2]. Indeed, pathogenic Alu insertions have been described in many cancer risk genes, such as *APC, ATM, CHEK2, BRCA1* and *BRCA2* [3]. In 2005, Teugels et al. described an Alu insertion within exon 3 of *BRCA2* (c.156_157insAlu) in a female breast cancer patient of Portuguese origin fulfilling clinical criteria of Hereditary Breast and Ovarian Cancer (HBOC) syndrome [4]. Subsequent analysis showed that the insertion co-segregated with the cancer phenotype in that family, and resulted in exon 3 skipping, a region that contains an important *BRCA2* regulatory domain [4]. Further, analysis has shown that *BRCA2* c.156_157insAlu is a Portuguese founder mutation accounting for 27–38% of all *BRCA1* and *BRCA2* pathogenic mutations in HBOC families originating mostly from northern/central Portugal [5,6]. Moreover, the estimated cumulative incidence of breast cancer in women carrying this mutation is at least as high as that among carriers of other *BRCA2* mutations [7].

Since Brazil was a Portuguese colony and received more than 2.2 million of Portuguese immigrants between 1500 and 1991 [8], in this study we sought to evaluate the frequency of this mutation in Brazilian individuals fulfilling the HBOC criteria. Furthermore, we performed ancestry and haplotype analyses of the mutation carriers to gain insight into the ancestral origin of the *BRCA2* c.156_157insAlu.

Materials and methods

Ethical aspects and patient recruitment

Unrelated individuals fulfilling HBOC criteria were recruited from three distinct Brazilian States, at three reference centers for Genetic Cancer Risk Assessment related to the Brazilian Hereditary Cancer Network (BHCN): Barretos Cancer Hospital (BCH, Barretos/São Paulo), Hospital de Clínicas de Porto Alegre (HCPA, Porto Alegre/Rio Grande do Sul) and Instituto Nacional de Câncer (INCA, Rio de Janeiro/Rio de Janeiro). Although all participants had a personal and/or family history suggestive of HBOC syndrome, each institution used slightly distinct testing criteria, detailed in Supplementary Methods and in Palmero et al. [9]. Of note, all patients recruited for this study fulfilled at least one of the NCCN HBOC genetic testing criteria. Ethical approval for this study was obtained from the institutional ethics committees of all participating centers and all patients were offered genetic counseling.

DNA isolation and *BRCA2* c.156_157insAlu genotyping

Genomic DNA was isolated from peripheral leucocytes using the Qiagen DNeasy kit or the Qiagen Flexigene kit following the manufacturer's instructions. *BRCA2* c.156_157insAlu genotyping was performed as previously described, using two distinct PCR reactions [5]. Positive and negative controls were used in all experiments and all positive cases were confirmed in a second independent sample.

Ancestry analyses

A panel of 46 AIM-INDELs was used and analyses were done in a single multiplex PCR followed by capillary electrophoresis, as previously described [10]. The electropherograms were analyzed and genotypes were automatically assigned with GeneMapper v4.1 (Applied Biosystems). Ancestry proportions were obtained using the Structure v2.3.3 software [11,12], taking into account the four main population groups: European (EUR), African (AFR), Native American (NAM) and East Asian (EAS). Ancestry was analyzed in the nine probands and 18 relatives with the *BRCA2* c.156_157insAlu variant.

Haplotype analyses

For haplotype analyses we assayed nine microsatellite markers flanking the *BRCA2* gene by PCR using fluorescently-labeled primers. PCR products were run on an ABI PRISM 310 Genetic Analyzer (Applied Biosystems) with the fluorescence labeled DNA fragment size standard TAMRA. Haplotype construction was performed manually based on the

Table 1 Detailed family history of the nine families carrying the *BRCA2* c.156_157insAlu.

Individual	Molecular and clinical characterization			Ancestry results			
	c.156_157insAlu status	Kinship	Diagnosis (age, years)	African (%)	European (%)	East Asian (%)	Amerindian (%)
INCA-1*	Carrier	Proband	BC (38)	3,3	93,6	2,2	0,9
INCA-1b	Non-carrier	Proband's mother	Unaffected	1,9	93,8	2,5	1,9
INCA-1c	Carrier	Proband's father	Gastric Ca (68)	5,3	90,9	2,7	1,1
INCA-2*	Carrier	Proband	BC (48)	28,7	64,3	4,4	2,6
INCA-2a	Non-carrier	Proband's niece	Unaffected	34,0	60,5	2,2	3,4
INCA-2b	Non-carrier	Proband's niece	Unaffected	27,5	67,0	3,3	2,2
INCA-2c	Carrier	Proband's daughter	Unaffected	17,0	64,1	2,3	16,5
INCA-2d	Carrier	Proband's sister	Unaffected	25,8	71,2	1,6	1,4
INCA-3*	Carrier	Proband	BC (33)	15,0	75,5	3,0	6,4
INCA-3a	Carrier	Proband's mother	BC (42)	14,2	70,0	5,2	10,7
HCPA-1	Carrier	Proband	Bilat BC (51, 55)	2,3	84,2	3,7	9,8
HCPA-2	Carrier	Proband	BC (27)	10,2	84,6	3,1	2,1
HCB −1*	Carrier	Proband	BC (39)	2,6	92,2	1,7	3,5
HCB −1a	Carrier	Proband's father	Unaffected	2,6	80,2	12,5	4,6
HCB-2*	Carrier	Proband	BC (39)	23,2	64,9	2,2	9,7
HCB-2a	Carrier	Proband's sister	Unaffected	4,2	90,7	2,8	2,3
HCB-2b	Carrier	Proband's sister	BC (62)	19,3	63,7	7,4	9,7
HCB-2c	Carrier	Proband's niece	BC (28)	4,6	85,8	6,2	3,4
HCB-2d	Carrier	Proband's sister	BC (62)	2,7	61,6	28,0	7,6
HCB-2e	Carrier	Proband's sister	Unaffected	4,6	91,4	2,3	1,7
HCB-2f	Carrier	Proband's son	Unaffected	21,5	37,7	16,8	24,1
HCB-2g	Carrier	Proband's brother	Unaffected	3,4	90,4	3,4	2,8
HCB-3*	Carrier	Proband	Gastric cancer (62), OC (62)	3,6	84,6	9,4	2,5
HCB-3a	Carrier	Proband's son	Unaffected	33,9	52,3	3,2	10,6
HCB-3b	Carrier	Proband's sister	Unaffected	19,2	52,9	1,7	26,1
HCB-4*	Carrier	Proband	BC (29)	4,0	84,0	5,0	5,0
HCB-4a	Carrier	Proband's father	Kidney Ca (47)	13,0	81,5	4,0	1,4

HCB: Hospital do Câncer de Barretos (São Paulo State); INCA: Instituto Nacional de Câncer (Rio de Janeiro State), HCPA: Hospital de Clínicas de Porto Alegre (Rio Grande do Sul State), BC: breast cancer, OC: ovarian cancer, Ca: cancer, Bilat: bilateral

genotypes obtained of probands and their relatives [5]. Haplotype analyses were performed in the nine probands and 18 relatives with *BRCA2* c.156_157insAlu variant.

Results

Screening for the *BRCA2* c.156_157insAlu was performed in 1380 apparently unrelated probands referred to Genetic Cancer Risk Assessment at Barretos (São Paulo State, $n=696$), Porto Alegre (Rio Grande do Sul State, $n=376$) and Rio de Janeiro (Rio de Janeiro State, $n=308$), due to a suspicion of HBOC syndrome. Nine carriers were identified, corresponding to an overall mutation prevalence of 0.65%. Four carriers were referred from São Paulo State, two from Rio Grande do Sul, and three from Rio de Janeiro, resulting in center-specific prevalences of 1.30%, 0.53% and 1.30%, respectively. Among carriers, 89% had breast cancer (8/9), with a mean age at diagnosis of 37.9 years. Only one proband had ovarian cancer, at 62 years old. Through genetic counseling and cascade testing of all available relatives, we identified 15 additional carriers in seven families.

Ancestry was analyzed in the nine probands and 18 relatives to estimate proportions of European (EUR), African (AFR), Amerindian (NAM) and East Asian (EAS) origin. All cases showed high proportion of European ancestry (range 38–94%), and the average ancestry proportions were 75.3% for EUR, 12.9% for AFR, 6.5% for NAM and 5.3% for EAS. Clinical, molecular and ancestry data are detailed in Table 1.

To evaluate the contribution of the founder Portuguese haplotype, all nine probands and their 18 relatives were genotyped for polymorphic microsatellite markers flanking *BRCA2*. Haplotypes were completely phased in three families (INCA-1, INCA-2 and HCB-2), differing from the ancestral haplotype only by one recombinational event (considering the most parsimonious relationships between these haplotypes). Additionally, the HCB-3 family, although not completely phased, also differed from the ancestral haplotype by one recombinational event. Although two other families (HCB-1 and HBC-4) were not completely phased, their haplotypes were compatible with the Portuguese ancestral haplotype (Table 2). In the remaining families (HCPA-1, HCPA-2 and INCA-3) it was not possible to infer the haplotype with the data available.

Table 2 Microsatellite markers genotypes in nine probands carrying the *BRCA2* c.156_157insAlu.

Markers	Consensus pattern	INCA-1	INCA-2	INCA-3	HCPA-1	HCPA-2	HCB-1	HCB-2	HCB-3	HBC-4	Portuguese ancestral haplotype H1
Centromere											
D13S1700	AGAA	**317**	253	**317**	249/305	329/333	**317**	249	**317**	**317**	**317**
D13S260	TG	**160**	162	**160**	154/168	160/162	**160**	162	**160**	**160**	**160**
D13S1698	GT	**156**	166	156/160	158/160	158/172	**156**	166	**156**	**156**	**156**
BRCA2											
D13S1701	TTCC	**299**	**299**	**299**	295/299	**299**	291/299	**299**	**299**	**299**	**299**
D13S171	TG	**230**	**230**	230/240	**230**	**230**	**230**	**230**	**230**	230/240	**230**
D13S1695	AC	**242**	**242**	258	242/252	242/250	**242**	**242**	252	242/250	**242**
D13S1694	TG	**228**	**228**	**228**	224/228	228/234	**228**	**228**	234	**228**	**228**
D13S310	GT	**144**	**144**	146	144/148	**144**	**144**	**144**	144/148	**144**	**144**
D13S267	TG	150	**158**	156	156/158	144/158	**158**	**158**	150	**158**	**158**
Telomere											

Markers present in the ancestral haplotype are highlighted in bold.

Discussion

Many current genetic testing methods applied to hereditary cancer cannot reliably detect large deletions and insertions related to retroelements, due to technical limitations [3]. In the past years, the two main diagnostic methods used in *BRCA* mutation screening were Sanger sequencing and Multiplex Ligation-dependent Probe Amplification (MLPA), and none of these methods were able to detect the c.156_156insAlu, which required a specific PCR-based test. While conventional PCR is still not capable of detecting such a large insertion as the *BRCA2* c.156_157insAlu, only recently (July 2016) a commercial MLPA kit for *BRCA2* was modified to include a specific probe for this mutation. Moreover, even considering the incorporation of next generation sequencing (NGS) into clinical genetic testing for hereditary cancer in the past few years, it is well known that the detection of large genomic rearrangements by NGS platforms is only possible after extensive optimization and validation, and requires a specific platform and bioinformatics pipeline which are not always available in commercial laboratories.

In this study, we found nine carriers of c.156_157insAlu among 1380 unrelated probands, corresponding to an overall mutation prevalence of 0.65%. From previous studies of Brazilian and non-Brazilian cohorts, we expected that nearly 20% of all individuals fulfilling standard (i.e., NCCN) criteria for HBOC would carry a pathogenic mutation in *BRCA1* or *BRCA2* [13–16]. Thus, in a cohort of 1380 individuals we would estimate the presence of approximately 276 *BRCA* mutation carriers. In this scenario, the nine proband carriers of the *BRCA2* c.156_157insAlu found in our study would represent around 3.3% of all detected mutations and about one tenth of the proportion of this mutation among all *BRCA* mutations observed in Portugal [5,6]. The lower mutational prevalence identified in our study is likely due to the high admixture observed in the Brazilian population, with genetic contributions from many distinct regions worldwide [17]. This is the consequence of five centuries of immigration and interethnic crosses of inhabitants from three areas: the European colonizers (mostly represented by the Portuguese), African slaves, and Amerindians. Although our ancestry analysis has shown a high proportion of European ancestry among all mutation carriers (75.3%), a recent study shows that individuals from the Southeast/South regions of Brazil derive from a wide European region, including central and northern Europe as well as parts of the Middle East. This profile is different from individuals originating in the northeast region of Brazil, which are expected to have a European ancestry mainly restricted to the Iberian Peninsula [18].

Among the nine proband carriers reported here, five were previously identified in smaller screening studies [13,19,20], and haplotype analysis was performed in three families without, however, a definitive conclusion about their ancestral haplotype [19]. Here, we performed haplotype reconstruction with nine microsatellites markers. Although the haplotype was not informative for three families, the remaining six presented haplotypes compatible with the ancestral Portuguese haplotype. These results suggest that these families share a common ancestor with the Portuguese *BRCA2* c.156_157insAlu positive families.

Despite the large number of Brazilian cases analyzed, our study is limited to the Southeast/South Brazilian regions. Thus, analysis in a larger cohort including individuals from all Brazilian regions would help to understand the overall contribution of *BRCA2* c.156_157insAlu in our population, especially in regions with high Portuguese ancestry. Moreover, the relatively low mutational prevalence found in this cohort does not rule out the importance of testing for this mutation in at-risk individuals and the absence of routine testing for this mutation in the past years may have led to an underestimation of its frequency in our population.

Although not as common as in Portugal, the c.156_157insAlu is the third most frequently reported *BRCA2* mutation in Brazil, corresponding to more than one third of all large genomic rearrangements reported in the Brazilian population (Palmero et al. Sci Rep. 2018 Jun 15;8(1):9188. doi:10.1038/s41598-018-27315-2), highlighting the importance of its identification.

Acknowledgments

We thank Gustavo Stumpf for his laboratory support and the Center of Molecular Diagnosis and the Oncogenetics Department from Barretos Cancer Hospital for their support.

Funding

This study was supported in part by grants from CNPq (408313/2016-1), FAPERGS (PRONEX: 16/2551-0000486-2), Hospital de Clínicas de Porto Alegre (FIPE-HCPA), FINEP-CT-INFRA (02/2010) and FAPESP (2013/24633-2). RMR, EIP and PA-P are recipients of National Council of Technological and Scientific Development (CNPq) scholarships.

Competing interests

The authors have no competing interests to declare.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cancergen.2018.09.001.

References

[1] Zhang W, Edwards A, Fan W, Deininger P, Zhang K. Alu distribution and mutation types of cancer genes. BMC Genomics 2011;12:157.
[2] Deininger PL, Batzer MA. Alu repeats and human disease. Mol Genet Metab 1999;67:183–93.
[3] Qian Y, Mancini-DiNardo D, Judkins T, Cox HC, Brown K, Elias M, et al. Identification of pathogenic retrotransposon insertions in cancer predisposition genes. Cancer Genet 2017;216-217:159–69.
[4] Teugels E, De Brakeleer S, Goelen G, Lissens W, Sermijn E, De Grève J. De novo Alu element insertions targeted to a sequence common to the BRCA1 and BRCA2 genes. Hum Mutat 2005;26:284.
[5] Peixoto A, Santos C, Rocha P, Pinheiro M, Príncipe S, Pereira D, et al. The c.156_157insAlu BRCA2 rearrangement accounts for more than one-fourth of deleterious BRCA mutations in northern/central Portugal. Breast Cancer Res Treat 2009;114:31–8.
[6] Peixoto A, Santos C, Pinheiro M, Pinto P, Soares MJ, Rocha P, et al. International distribution and age estimation of the Portuguese BRCA2 c.156_157insAlu founder mutation. Breast Cancer Res Treat 2011;127:671–9.
[7] Mavaddat N, Peock S, Frost D, Ellis S, Platte R, Fineberg E, et al. Cancer risks for BRCA1 and BRCA2 mutation carriers: results from prospective analysis of EMBRACE. J Natl Cancer Inst 2013;105:812–22.
[8] IBGE. Brasil: 500 anos de povoamento. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística (IBGE); 2007.
[9] Palmero EI, Galvão HC, Fernandes GC, Paula AE, Oliveira JC, Souza CP, et al. Oncogenetics service and the Brazilian public health system: the experience of a reference cancer hospital. Genet Mol Biol 2016;39:168–77.
[10] Pereira R, Phillips C, Pinto N, Santos C, dos Santos SE, Amorim A, et al. Straightforward inference of ancestry and admixture proportions through ancestry-informative insertion deletion multiplexing. PLoS One 2012;7:e29684.
[11] Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics 2000;155:945–59.
[12] Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 2003;164:1567–87.
[13] Fernandes GC, Michelli RA, Galvão HC, Paula AE, Pereira R, Andrade CE, et al. Prevalence of BRCA1/BRCA2 mutations in a Brazilian population sample at-risk for hereditary breast cancer and characterization of its genetic ancestry. Oncotarget, 2016:80465–81.
[14] Kast K, Rhiem K, Wappenschmidt B, Hahnen E, Hauke J, Bluemcke B, et al. Prevalence of BRCA1/2 germline mutations in 21 401 families with breast and ovarian cancer. J Med Genet 2016;53:465–71.
[15] Alemar B, Gregório C, Herzog J, Matzenbacher Bittar C, Brinckmann Oliveira Netto C, Artigalas O, et al. BRCA1 and BRCA2 mutational profile and prevalence in hereditary breast and ovarian cancer (HBOC) probands from Southern Brazil: are international testing criteria appropriate for this specific population? PLoS One 2017;12:e0187630.
[16] Alemar B, Herzog J, Brinckmann Oliveira Netto C, Artigalás O, Schwartz IV, Matzenbacher Bittar C, et al. Prevalence of hispanic BRCA1 and BRCA2 mutations among hereditary breast and ovarian cancer patients from Brazil reveals differences among Latin American populations. Cancer Genet 2016;209:417–22.
[17] Ruiz-Linares A, Adhikari K, Acuña-Alonzo V, Quinto-Sanchez M, Jaramillo C, Arias W, et al. Admixture in Latin America: geographic structure, phenotypic diversity and self-perception of ancestry based on 7,342 individuals. PLoS Genet 2014;10:e1004572.
[18] Kehdy FS, Gouveia MH, Machado M, Magalhães WC, Horimoto AR, Horta BL, et al. Origin and dynamics of admixture in Brazilians and its effect on the pattern of deleterious mutations. Proc Natl Acad Sci USA 2015;112:8696–701.
[19] Ewald IP, Cossio SL, Palmero EI, Pinheiro M, Nascimento IL, Machado TM, et al. BRCA1 and BRCA2 rearrangements in Brazilian individuals with hereditary breast and ovarian cancer syndrome. Genet Mol Biol 2016;39:223–31.
[20] Moreira MA, Bobrovnitchaia IG, Lima MA, Santos AC, Ramos JP, Souza KR, et al. Portuguese c.156_157insAlu BRCA2 founder mutation: gastrointestinal and tongue neoplasias may be part of the phenotype. Fam Cancer 2012;11:657–660.

ANEXO II – Arquivo PDF do artigo publicado na revista científica "Oncotarget".

www.oncotarget.com Oncotarget, 2018, Vol. 9, (No. 44), pp: 27525-27534

Research Paper

Genetic alterations detected by comparative genomic hybridization in BRCAX breast and ovarian cancers of Brazilian population

Paula Silva Felicio[1], Lucas Tadeu Bidinotto[1,2], Matias Eliseo Melendez[1], Rebeca Silveira Grasel[1], Natalia Campacci[1], Henrique C.R. Galvão[3], Cristovam Scapulatempo-Neto[4], Rozany Mucha Dufloth[4], Adriane Feijó Evangelista[1] and Edenir Inêz Palmero[1,2]

[1]Molecular Oncology Research Center, Barretos Cancer Hospital, Barretos, SP, Brazil

[2]Barretos School of Health Sciences-FACISB, Barretos, SP, Brazil

[3]Department of Oncogenetics, Barretos Cancer Hospital, Barretos, SP, Brazil

[4]Department of Pathology, Barretos Cancer Hospital, Barretos, SP, Brazil

Correspondence to: *Edenir Inêz Palmero,* ***email:*** *edenirip@yahoo.com.br*

Keywords: *BRCAX; hereditary breast cancer; hereditary breast and ovarian cancer predisposition syndrome; CNV alterations*

Received: *August 09, 2017* ***Accepted:*** *May 14, 2018* ***Published:*** *June 08, 2018*

ABSTRACT

Background: About 5–10% of breast/ovarian cancers are hereditary. However, for a large proportion of cases (around 50%), the genetic cause remains unknown. These cases are grouped in a separated BRCAX category. The aim of this study was to identify genomic alterations in *BRCA1/BRCA2* wild-type tumor samples from women with family history strongly suggestive of hereditary breast/ovarian cancer.

Results: A cohort of 31 Brazilian women was included in the study. Using the GISTIC algorithm, we identified 20 regions with genomic gains and 31 with losses. The most frequent altered regions were 1q21.2, 6p22.1 and 8p23.3 in breast tumors and Xq26 and Xp22.32-22.31 among the ovarian cancer cases. An interesting association identified was the loss of 22q13.31-13.32 and the presence of ovarian cancer cases. Among the genes present in the frequently altered regions, we found *FGFR1*, *NSMCE2*, *CTTN*, *CRLF2*, *ERBB2*, *STARD3*, *MIR3201* and several genes of *RAET* and *ULBP* family.

Conclusions: In conclusion, our results suggest that alterations on chromosomes 1, 6, 8 and X are common on BRCAX tumors and that the loss on 22q can be associated with the presence of ovarian cancer.

Methods: DNA copy number alterations were analyzed by 60K array comparative genomic hybridization in breast and ovarian FFPE tumors.

INTRODUCTION

According to the World Health Organization, breast cancer (BC) is the most common tumor in women worldwide [1]. It is known that 5–10% of BC cases have a hereditary component [2], being characterized by the presence of germline mutations in the *BRCA1* [3] or *BRCA2* [4] genes, which are associated with the hereditary breast and ovarian cancer predisposition syndrome (HBOC). HBOC patients have strong personal and family histories of cancer. Moreover, these patients are characterized by early age-at-diagnosis of cancer, increased frequency of bilateral tumors, and two or more generations affected by cancer [5, 6].

Recent studies have shown that alterations in other susceptibility genes, mainly involved in the homologous recombination and DNA repair pathways, can be causal factors of hereditary breast and ovarian cancers [7]. In spite of that, the predisposing genetic cause of about 50% of the families at-risk for hereditary breast and ovarian cancers remains unknown [8, 9]. These families are grouped in a category called BRCAX.

Evidences from the literature have shown that BRCAX tumors are rather heterogeneous, involving

several different histopathological subgroups and genetic alterations [10, 11]. Several authors have shown the presence of new high penetrance genes associated with breast and ovarian cancers [11–17]. However, the opinion of the scientific community is controversial. There are authors who argue that the incidence of BRCAX tumors is associated with rare syndromes in which BC is only one component [12, 15, 16]. Other authors believe that this type of tumor results from mutations in several genes with low penetrance or population-specific [11, 13, 14, 18].

Studies using array-comparative genomic hybridization (aCGH) technique suggest that several chromosomal regions are associated with the development of hereditary BC, highlighting gains at chromosomes 1q, 8q, 17q and 20q and losses within chromosomes 8p, 11q, 13q and 17p [19–24]. Despite these findings, more studies are necessary to a better understanding of BRCAX molecular events in hereditary breast cancer. In this regard, the aim of this study was to identify chromosomal and subchromosomal copy number alterations in tumor samples from Brazilian women without *BRCA1/BRCA2* germline mutations with family history strongly suggestive of HBOC syndrome.

RESULTS

In the present study, we analyzed 31 Brazilian women at-risk for hereditary breast/ovarian cancer (27 with personal history of BC and 4 with ovarian tumors) without *BRCA1/BRCA2/TP53* germline mutations, by array-CGH. Clinicopathological characteristics and family history of the patients are specified in Table 1.

Briefly, the mean age at BC diagnosis was 42.9 years (SD = 7.9), ranging from 27–70 years. The majority of BC was invasive ductal carcinoma (77.8%), estrogen and progesterone positive (69.2% and 73.1%, respectively) and HER2 negative (64.0%). Regarding molecular classification, the majority of patients presented luminal type tumors (21 patients, 80.8%), four patients (15.4%) had triple negative tumors and only one patient (3.8%) was diagnosed with a HER2 subtype tumor.

All four ovarian cancer patients developed serous adenocarcinoma subtype tumors. The average age at diagnosis was 47.7 years (SD = 18.0), ranging from 21–60 years.

A detailed cancer family history can be found in Table 1. All patients reported at least one case of BC in the family, diagnosed at early age (<55 years for breast cancer cases). In addition, two women at-risk for hereditary BC (samples: 960 and 1024) had a family history with bilateral BC. Among patients diagnosed with BC, the majority reported more than three cases of BC in the family history (16 cases, 59.3%). Meanwhile, all patients diagnosed with ovarian cancer, reported three or less BC cases in their families ($p = 0.043$). Moreover, 12 patients reported the presence of ovarian cancer in the family history.

The molecular analysis revealed gained and lost regions across all chromosomes for both breast and ovarian tumors (Figure 1). We found 20 gained regions and 31 lost in BRCAX tumors. In addition, some variations, although not statistically significant, were found only in patients diagnosed with breast cancer, such as: gains of 7p22.1, 12p13.1, 14q13.3-q21.1, 17q11.2, 17q12 and 17q21.32-q21.33, and losses of 2p25.3, 6q25.3-q26 and 10q26.3. Moreover, the gain of Xq26 and loss of Xp22.32–22.31 were more frequent in ovarian cancer (100%), compared with breast cases (26% and 59%, respectively) ($p = 0.01$ for both regions). Loss of 22q13.31–13.32 was detected more often in ovarian than in breast cancer cases ($p = 0.043$). In addition, a significant number of copy number alterations involving chromosome 8 was observed.

When family history was taken into consideration for copy number variation analyses, we observed that loss of 22q13.31–13.32 region was significantly associated with the presence of ovarian family history ($p = 0.03$). This region includes *MIR3201*, *LOC284933*, *FAM19A5*, *MIR4535*, *LINC01310* genes. Other significant association found included gains in the 6p22.1 region (including 13 histone family genes) in 100% of metastatic cases ($p = 0.03$). Finally, we found loss of 6q25.1 in 71% of patients with metastasis ($p = 0.01$). This region includes *RAET1E*, *RAET1E-AS1*, *RAET1G*, *ULBP2*, *ULBP1*, *RAET1K*, *RAET1L*, *ULBP3*, *PPP1R14C*, *IYD*, *PLEKHG1*, *MTHFD1L* genes (Supplementary Table 1).

In addition, when comparing our findings with those of the literature of BRCAX tumors, we observed that our results corroborate some findings reported by Didraga *et al.* (2011), Alvarez *et al.* (2016) and Mangia *et al.* (2008), showing 50%, 21% and 12% of common regions, respectively (Figure 2).

Finally, we found that 22 genes present in gained regions also present overexpression in the Oncomine database, whereas 21 genes present in lost regions show loss of expression in the same database ($p < 0.01$, Table 2).

DISCUSSION

In the present study, a BRCAX tumor characterization of FFPE samples has been performed by array comparative genomic hybridization. Among the altered loci, we can highlight the identification of several alterations in chromosome 8, including losses on 8p12-p11.23 and gains on 8p12-p11.23 and 8q24.13, in concordance with previous studies of BRCAX tumors [14, 25, 26]. Besides, the chromosomal region 8p12-p11 has been reported to be amplified in 10–23% of BC cases [27–29], and some studies have shown that amplification on this region is associated with poor clinical outcome [27, 30]. We found by *in silico* analysis that 4 genes present in this region (including *FGFR1* and *NSMCE2)* are overexpressed.

The *FGFR1* gene encodes a transmembrane protein that interacts with fibroblast growth factors and directly influence mitogenesis and cell differentiation. In fact, there are several studies showing different treatment outcomes

Table 1: Clinicopathological characteristics and family history of the patients at-risk for hereditary cancer

Family	Cancer (age at diagnosis)	Histological type	Molecular subtype	Breast/Ovarian cancer cases in the family (sex and age at diagnosis, if known)
19	Breast (44)	IDC	ER: –; PR: +; HER2: –	Sister: Breast (F,46; F,46)
29	Ovarian (42), Breast (53)	DCIS	ER: +; PR: +	Paternal side of the family: Breast (F,29); Ovarian (F,60; F,?; F, ?), Uterus (F,57; F,?; F,?); Gastric (M,42; M,?; M,?; M,?)
65	Breast (35)	IDC	ER: –; PR: –; HER2: –	Maternal side of the family: Breast (F,31; F,34; F,47; F,39; F,39; F,46)
80	Breast (43)	DCIS	ER: +; PR: +; HER2: +	Maternal side of the family: Breast (F,44; f,44; F,55; F,57; F,60), Prostate (M,?)
85	Breast (51)	IDC	ER: +; PR:+; HER2: –	Maternal side of the family: Breast (F,43; F,45; F,48), Stomach (F;45; M,56); Leukemia (M,69)
179	Breast (47)	IDC	ER: +; PR: +	Paternal side of the family: Breast (F,37; F,49; F,61), Throat (M,?; M,?)
233	Breast (49)	IDC	ER: –; PR: –; HER2: –	Maternal side of the family: Breast (F,?; F,50; F,33; F,70; F,60; F,60; F,46), Colorectal (M,65), Gastric (M,62), Pancreas (M,62), Lung (M,52; M,66; M,?)
241	Breast (45)	IDC	ER: +; PR: +; HER2: –	Paternal side of the family: Breast (F,48; F,49)
275	Ovarian (60)	Serous adenocarcinoma	Not applicable	Paternal side of the family: Breast (F,32; F,35), Prostate (M,80)
289	Breast (48)	IDC	ER: –; PR: –; HER2: +	Maternal side of the family: Breast (F,50; F,65; F,65)
306	Melanoma (26), Breast (36)	IDC	ER: +; PR: +; HER2: –	Maternal side of the family: Breast (F,43; F,?)
320	Ovarian (53)	Serous adenocarcinoma	Not applicable	Maternal side of the family: Breast (F,52), Ovarian (F,71), Uterus (F,60), Thyroid (M,29), Lung (M,83)
426	Breast (38)	IDC	ER: +; PR: +; HER2: –	Paternal side of the family: Breast (F,20); Ovarian (F,28), Leukemia (M,78), Esophagus (M,?)
494	Breast (33)	IDC	ER: +; PR: +; HER2: +	Maternal side of the family: Breast (F,38; F,?), Ovarian (F,38)
558	Breast (37)	IDC	ER: –; PR: –; HER2: –	Maternal side of the family: Breast (F,52; F,?; F,?; F,?, F,?), Ovarian (F,42), Skin (F,?)
563	Breast (39)	IDC	ER: +; PR: –; HER2: –	Paternal side of the family: Breast (F,30; F,40; F,45, F,50; F,51), Lung, (M,?), Colorectal (F,64), Skin (M,72)
581	Breast (45)	DCIS	Not available	Paternal side of the family: Breast (F,49; F,46; F,54), Prostate (M,60; M,70), head and neck (M,83)
593	Breast (39)	IDC	ER: +; PR: +; HER2: –	Paternal side of the family: Breast (F,50; F,?; F,?; F,?), Ovarian (F,?; F,?), Gastric (M,?; M,?), Colorectal (M,?; M,?)
626	Breast (46)	DCIS	ER: +; PR: +	Maternal side of the family: Breast (F,74; F,80; F,57; F,45), Ovarian (F,45), Thyroid (F,40), Skin (M,80), Pancreas (M,?), Myeloma (M,60), Lips (M,?)
638	Breast (42)	ILC	ER: +; PR: +	Maternal side of the family: Breast (F,49; F,50; F,?), Gastric (F,55), Thyroid (F,36), Lips (F,55)
649	Breast (38)	IDC	ER: +; PR: +; HER2: –	Maternal side of the family: Breast (F,64), Ovarian (F,61), Thyroid (F,61)
695	Ovarian (21)	Serous adenocarcinoma	Not applicable	Paternal side of the family: Breast (F,42), Ovarian (F,68), Colorectal (F,40; M,40), Gastric (F,50; F,70)

960	Bilateral Breast (59,70)	IDC	ER: +; PR: +; HER2: –	Maternal side of the family: Breast (F,34; F,59), Uterus (F,45), Lung (M,77; M,?)
974	Breast (46)	IDC	ER: +; PR: +; HER2: –	Maternal side of the family: Breast (F,55; F,45; F,60; F,60; F,55; F,45; F,60), Prostate (M,70; M,80)
981	Breast (37)	IDC	ER: +; PR: +; HER2: –	Maternal side of the family: Breast (F,32; F,70; F,60), Melanoma (F,30; F36), Leukemia (F,5), Bile ducts (M,49; F,55)
1014	Breast (42)	DCIS	ER: +; PR: +	Maternal side of the family: Breast (F,53; F,?); Melanoma (F,75), Lymphoma (M,19), Liver (F,?), Brain (F,?)
1024	Breast (48)	IDC	ER: –; PR: –; HER2: –	Paternal side of the family: Breast (F,70; F,72; F,44; F;44; F,49), Ovarian (F,56), Colorectal (M,20), Melanoma (M, ?), Prostate (M,50), Gastric (F,70; F,72; F,41)
1055	Ovarian (57)	Serous adenocarcinoma	Not applicable	Maternal side of the family: Breast (F,49; F,50), Pancreas (F,50), Lung (M,?)
1095	Breast (43)	IDC	ER: –; PR: –; HER2: +	Paternal side of the family: Breast (F,27; F,42), Uterus (F,98), Throat (M,72)
1151	Breast (38)	IDC	ER: +; PR: +; HER2: –	Maternal side of the family: Breast (F,35; F,60)
1264	Breast (27)	IDC	ER: –; PR: +; HER2: –	Maternal side of the family: Breast (F,50), Pancreas (M,75); Intestine (M,81)

of breast cancer women depending on the *FGFR1* status [31–33]. Similarly, *NSMCE2* plays an important role in cell cycle, since its depletion in MCF-7 breast cancer cells affected cell cycle and G1-S transition [34]. Moreover, the overexpression of cortactin (*CTTN*), present in 11q13.3, was linked to *CCND1* amplification in premenopausal breast cancer [35], although it failed to demonstrate a strong prognostic value in patients with breast cancer [36]. Conversely, its upregulation promoted colon cancer progression through ERK pathway [37]. Therefore, other studies have shown that amplification on chromosomal region 8p12-p11 in combination with amplification

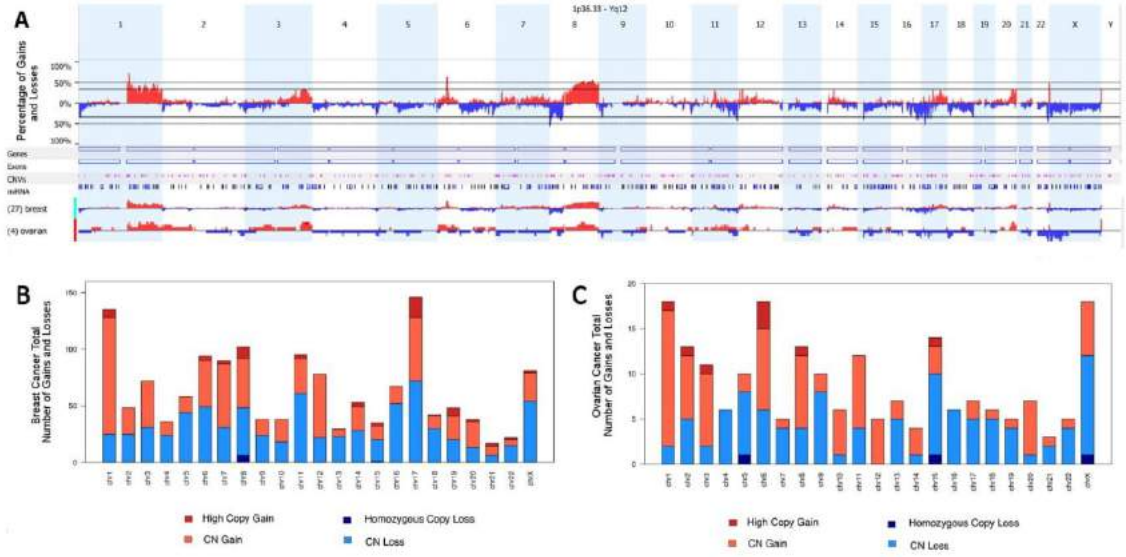


Figure 1: Overview of gained and lost regions across all chromosomes. (**A**) Overall and specific breast and ovarian copy number aberration frequencies. Regions presenting copy gains are shown in red and with copy loss in blue. (**B**) Overview of gained and lost regions across all chromosomes in breast tumors. (**C**) Overview of gained and lost regions across all chromosomes in ovarian tumors.

on 11q13 have more impact on patient outcome than amplification on only one of the two loci [27, 38].

In addition to gains and losses on chromosome 8 and alterations on chromosome 11, alterations in chromosome X seem to be characteristic of BRCAX tumors. In our study, a great number of samples showed gains on regions 11q13.2-q13.3 and Xp22.33, which were also identified by Didraga and collaborators [25]. Although it is not extensively studied in breast cancer, the overexpression of *CRLF2*, present in Xp22.33, has been demonstrated to be a marker of poor outcome of pediatric and adult B-precursor acute lymphoblastic leucemia (ALL) (as reviewed in [39]).

Study performed by Gronwald *et al.* [19] compared BRCAX with sporadic breast cancers and identified

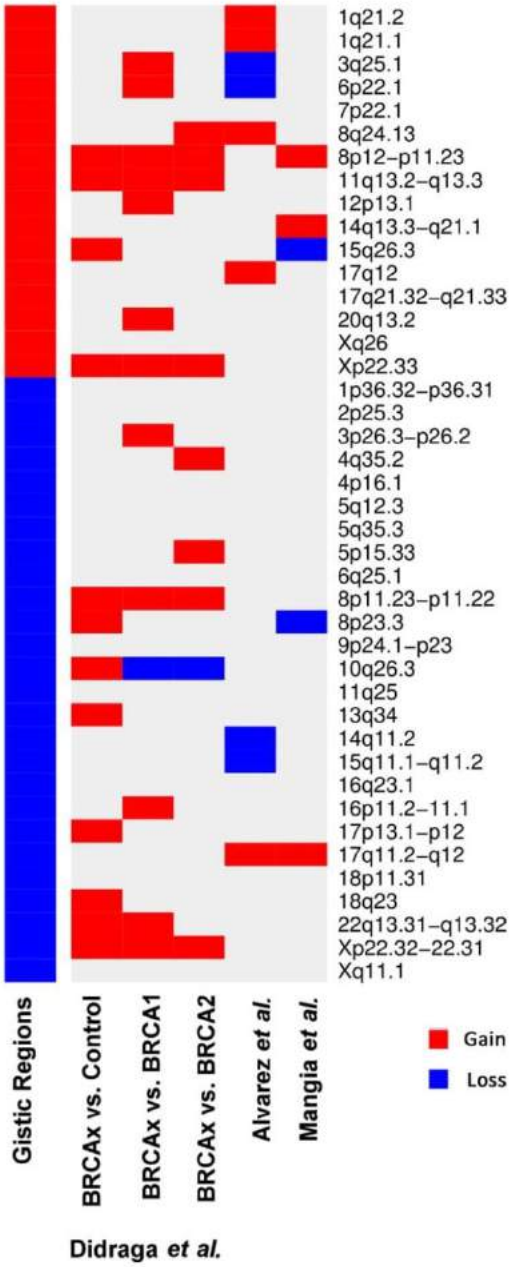


Figure 2: Heatmap representing the gains (in red) and losses (in blue) through aCGH found by GISTIC algorighm in common with previous studies by Didraga *et al.* (2011), Alvarez *et al.* (2016) and Mangia *et al.* (2008).

Table 2: Genes in gained regions that presented *in silico* overexpression and genes in lost regions that presented *in silico* loss of expression

Event[1]	Cytoband	Genes
Gain/Overexp	1q21.1-q21.2	*PEX11B, PDE4DIP, ECM1, TARS2, RPRD2*
Gain/Overexp	6p22.1	*HIST1H3H, HIST1H4J, HIST1H4K*
Gain/Overexp	7p21.1	*HDAC9*
Gain/Overexp	8p11.23-p11.22	*TM2D2, LETM2, RNF5P1, FGFR1*
Gain/Overexp	8q24.13	*NSMCE2, KIAA0196*
Gain/Overexp	11q13.3	*CTTN*
Gain/Overexp	17q12	*ERBB2, STARD3, GRB7*
Gain/Overexp	17q21.32-q21.33	*PHB, ABI3*
Gain/Overexp	Xp22.33	*CRLF2*
Loss/LOExp	1p36.32	*TPRG1L, AJAP1*
Loss/LOExp	2p25.3	*FAM150B, TMEM18, TPO*
Loss/LOExp	3p26.3-p26.2	*CRBN, CNTN4*
Loss/LOExp	5q35.3	*ADAMTS2, ZNF879, COL23A1*
Loss/LOExp	8p23.3	*ERICH1, RPL23AP53, OR4F21, ZNF596*
Loss/LOExp	11q25	*JAM3, LOC283177, THYN1*
Loss/LOExp	14q11.2	*OR4K5*
Loss/LOExp	16q23.1	*CNTNAP4, SYCE1L*
Loss/LOExp	Xp22.32-p22.31	*NLGN4X*

[1]Events represent Gain or Loss in our samples with concurrent Overexpression (Overexp) or Loss of Expression (LOExp) on Oncomine samples.

several altered regions (114 gains and 36 losses) in 18 patients. Their findings showed concordances with our results, presenting more often gains in 1q, 6p, 17q and frequent loss of 8p. Beside the well known effects of *ERBB2* amplification in breast cancer development, the overexpression of *STARD3* (located in the same locus) seems to be important, since it may contribute to increased proliferation, migration and invasion of breast cancer cells (as reviewed in [40]). Finally, considering our findings of altered regions found in BRCAX associated with metastasis (gain of 6p22.1 and loss of 6q25.1), gain of 6p was previously associated with BRCAX, and loss of 6q with *BRCA1* tumors [41]. In fact, there are several members of *RAET* and *ULBP* family present in this locus. These members are ligands of C-type lectin-like receptor NKG2D, present in NK and T cells subsets, highly involved in tumor immunosurveillance [42]. Therefore, the loss of this region may have led to lower expression of these ligands, leading to less immunogenicity of the tumor cells. In fact, there are reports on colorectal cancer that have demonstrated this same pattern, and several authors discuss the potential therapeutic utility of NKG2D ligands in the treatment of this disease [42–44]. Therefore, these alterations on chromosome 6 seem to be highly associated with breast cancer tumors and may be of interest for further studies.

We also found that loss of 22q13.31–13.32 was significantly associated with presence of ovarian tumors (in the proband or in the family). The loss of heterozygosis (LOH) of chromosome 22q has been reported in a variety of cancers, including ovarian cancers, where the LOH rates reached 70% of cases [45, 46]. Study published by Zweemer *el al.* (2001) reported a significant loss of 22q, identified through aCGH in familial ovarian tumors [47]. Interestingly, *MIR3201* was significantly downregulated in recurrent epithelial ovarian cancer (EOC), when compared to primary EOC [48]. To the best of our knowledge, there are no studies pointing to the functional relevance of *MIR3201* in ovarian cancer, however, further studies may be performed to evaluate its possible role as a biomarker of EOC recurrence.

In summary, our findings support previous data of BRCAX related alterations and point to new regions potentially associated with personal and family history of ovarian cancer. In the present study, we could identify by aCGH analysis a potentially BRCAX-associated ovarian region on chromosome 22. Given our limited sample size, further work should be performed in order to validate our findings, to identify the driver genes associated with the BRCAX tumor development, as well as to uncover the role of those altered regions in cancer formation and progression.

MATERIALS AND METHODS

Ethics statement

All participants gave their consent to use tumor samples for academic genetic research. In addition, the ethics committee of the Barretos Cancer Hospital (BCH) approved this study (approval number: 916/2015).

Patients

This study included 31 unrelated Brazilian women at-risk for hereditary breast and ovarian cancer from the Oncogenetics Department of BCH. Those women were referred from the Oncogenetics Department of BCH for *BRCA1, BRCA2* and *TP53* genetic testing due to the presence of clinical criteria for HBOC, but no genetic alterations in these genes were found. For the purpose of the present study, were included only families fulfilling the following criteria: patients diagnosed with breast/ovarian cancer at an early age (<55 years), with at least two relatives with breast and/or ovarian cancer, two or more generations affected by cancer and absence of male BC.

Clinical information was obtained through detailed review of the patient´s clinical chart. For family history data, all pedigrees were revised.

Sequencing of *BRCA1, BRCA2* and *TP53*

Analysis of the presence of germline mutations in *BRCA1/BRCA2/TP53* genes was conducted at the Center of Molecular Diagnosis of BCH as part of routine care through NGS sequencing followed by rearrangement analysis through MLPA (Multiplex Ligation-dependent Probe Amplification Analysis), as described elsewhere by Fernandes *et al*. [49].

Tumor samples

For aCGH analysis, a representative section of FFPE tumor tissue from the breast or the ovarian tumor was stained by hematoxylin and eosin (H&E) and evaluated by a pathologist to verify tumor content (>70% tumor) and further microdissection.

DNA isolation and quality control

Following microdissection, DNA extraction steps were carried out using DNeasy Blood and Tissue kit (*Qiagen*), following the manufacturer's instructions. The quality and integrity of the extracted DNA was assessed by multiplex PCR reaction using four primer pairs for the *GAPDH* gene (amplifying 100, 200, 300 and 400 bp, respectively), as described by Van Beers *et al*. [50]. The PCR reaction carried out contained (in a final volume of 30 µL) 1.5 mM $MgCl_2$; 0.2 mM dNTP (*Invitrogen*); 0.133 µM of each primer; 1 U Taq DNA polymerase (*Invitrogen*) and 60 ng of tumor DNA. Reactions were performed in a *Veriti* thermocycler (*Thermo Fisher Scientific*) using the following amplification parameters: 94° C for 1 minute, 35 cycles of 94° C for 1 minute, 56° C for 1 minute, and 72° C for 3 minutes. Finally, a final extension at 72° C for 7 minutes. Amplification of DNA was verified by agarose gel electrophoresis.

Array comparative genomic hybridization

aCGH was performed on oligonucleotide-based SurePrint G3 Unrestricted CGH 8 × 60 K microarray slides, according the protocol provided by the manufacturer. In brief, 1 µg in final volume of 13 µL of normal female control DNA – reference DNA (DNA universal control-Promega Madison WI USA- Woman Reference: G152A) and patient's DNA were differentially labeled with Cy3 (cyanine 3-deoxyuridine triphosphate) and Cy5 (cyanine 5-deoxyuridina triphosphate), respectively, using Agilent SureTag Complete DNA Labeling Kit (*Agilent Technologies*). Labeled DNA was then cleaned with purification columns (*Agilent Technologies*) and hybridized on array at 65° C for 24 hours, according to manufacturer's recommendations. Microarrays were washed using Agilent Oligo aCGH Wash Buffers and scanning was performed using Agilent SureScan Microarray Scanner according to manufacturer's instructions (*Agilent Technologies*).

Data analysis

Data quantification of aCGH was performed with Feature Extraction software (*Agilent Technologies*) and the txt output files were imported into Nexus Copy Number v8.0 (*BioDiscovery Inc*) for visualization and downstream analysis. BioDiscovery's FASST2 Segmentation Algorithm, a Hidden Markov Model (HMM) based approach, was used to make copy number calls. The FASST2 algorithm, unlike other common HMM methods for copy number estimation, does not aim to estimate the copy number state at each probe but uses many states to cover more possibilities, such as mosaic events. These state values are then used to make calls based on a log-ratio threshold. The significance threshold for segmentation was set at 5.0E-6 also requiring a minimum of 3 probes per segment and a maximum probe spacing of 1,000 kb between adjacent probes before breaking a segment. The log ratio thresholds for single copy gain (or amplification) and single copy loss (or deletion) were set at 0.2 and -0.23, respectively. The log ratio thresholds for two or more copy gain (or high copy gain) and homozygous loss (or high copy loss) were set at 1.14 and -1.1 respectively. A 3:1 sex chromosome gain threshold was set to 1.2 and a 4:1 sex chromosome gain threshold was set to 1.7. Male sex chromosome big loss threshold was set to -1.1. GISTIC (Genomic Identification of Significant Targets in Cancer)

algorithm was used within Nexus 8.0 to identify regions that are significantly amplified or deleted across a set of samples. It was considered the default parameters of Q-bound ≤ 0.05 with False Discovery Rate (FDR) correction and G-score cut-off ≤ 1.0. The identification of genes and CNVs were also performed within Nexus 8.0, being CNVs filtered according to 1000 genomes project. It was calculated the frequency of the gained and lost remaining CNVs and further separated according to <1% (rare CNVs) and ≥1% (common CNVs). The peaks identified by GISTIC algorithm were associated to breast and ovarian cancer family history and clinical characteristics, i.e. clinical staging, age at diagnosis (≤30, 31–45 and ≥45 years), molecular subtype, histological subtype, presence of metastasis and recurrence. These analyses were done by Fisher's exact test (within SPSS v.21.0 software for Windows (Chicago, IL) considering the significance level of 5%.

Besides, the genomic regions found to be significant in GISTIC were considered for further analysis using the professional version of the compendium of cancer transcriptome profiles, Oncomine™ (Compendia Bioscience, Ann Arbor, MI). There were selected 13 breast and 5 ovarian cancer datasets (totalizing over 4000 samples). For each cancer type (breast or ovary), we selected the genes that presented gain or loss in our aCGH, and considered relevant those that presented gain in our aCGH and overexpression in Oncomine ($P < 0.01$), or those that presented loss in our aCGH and loss of expression in Oncomine ($P < 0.01$).

Abbreviations

aCGH: Array comparative genomic hybridization; BC: breast cancer; BCH: Barretos Cancer Hospital; ER: estrogen receptor; HBOC: Hereditary Breast and Ovarian Cancer Predisposition Syndrome; HER2: Human Epidermal growth factor Receptor 2; PR: progesterone receptor.

ACKNOWLEDGMENTS

The authors thank the patients, the Pathology Department and the Center of Molecular Diagnosis of Barretos Cancer Hospital.

CONFLICTS OF INTEREST

The authors declare that they have no competing interests.

FUNDING

This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP - 2013/24633-2). EIP is recipient of a National Council of Technological and Scientific Development (CNPq) scholarship.

REFERENCES

1. WHO. World Health Organization - Breast Cancer. 2016. available from: http://www.who.int/cancer/detection/breastcancer/en/.
2. Lynch HT, Snyder C, Lynch J. Hereditary breast cancer: practical pursuit for clinical translation. Ann Surg Oncol. 2012; 19:1723–31. https://doi.org/10.1245/s10434-012-2256-z.
3. Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, Cochran C, Bennett LM, Ding W, Bell R, Rosenthal J, Hussey C, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science. 1994; 266:66–71.
4. Wooster R, Neuhausen SL, Mangion J, Quirk Y, Ford D, Collins N, Nguyen K, Seal S, Tran T, Averill D. Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. Science. 1994; 265:2088–90.
5. Gronwald J, Byrski T, Huzarski T, Oszurek O, Janicka A, Szymanska-Pasternak J, Gorski B, Menkiszak J, RzepkaGorska I, Lubinski J. Hereditary breast and ovarian cancer. Hered Cancer Clin Pract. 2008; 6:88–98. https://doi.org/10.1186/1897-4287-6-2-88.
6. Petrucelli N, Daly MB, Pal T. BRCA1- and BRCA2-Associated Hereditary Breast and Ovarian Cancer. In: Pagon RA, Adam MP, Ardinger HH, Wallace SE, Amemiya A, Bean LJH, Bird TD, Ledbetter N, Mefford HC, Smith RJH and Stephens K, eds. GeneReviews(R). (Seattle (WA). 1993. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1247/.
7. Friedenson B. BRCA1 and BRCA2 pathways and the risk of cancers other than breast or ovarian. MedGenMed. 2005; 7:60.
8. Couch FJ, Nathanson KL, Offit K. Two decades after BRCA: setting paradigms in personalized cancer care and prevention. Science. 2014; 343:1466–70. https://doi.org/10.1126/science.1251827.
9. Mavaddat N, Antoniou AC, Easton DF, Garcia-Closas M. Genetic susceptibility to breast cancer. Mol Oncol. 2010; 4:174–91. https://doi.org/10.1016/j.molonc.2010.04.011.
10. Tanic M, Andres E, Rodriguez-Pinilla SM, Marquez-Rodas I, Cebollero-Presmanes M, Fernandez V, Osorio A, Benitez J, Martinez-Delgado B. MicroRNA-based molecular classification of non-BRCA1/2 hereditary breast tumours. Br J Cancer. 2013; 109:2724–34. https://doi.org/10.1038/bjc.2013.612.
11. Honrado E, Osorio A, Milne RL, Paz MF, Melchor L, Cascon A, Urioste M, Cazorla A, Diez O, Lerma E, Esteller M, Palacios J, Benitez J. Immunohistochemical classification of non-BRCA1/2 tumors identifies different groups that demonstrate the heterogeneity of BRCAX

families. Mod Pathol. 2007; 20:1298–306. https://doi.org/10.1038/modpathol.3800969.

12. Hopper JL. More breast cancer genes? Breast Cancer Res. 2001; 3:154–7.

13. Oldenburg RA, Meijers-Heijboer H, Cornelisse CJ, Devilee P. Genetic susceptibility for breast cancer: how many more genes to be found? Crit Rev Oncol Hematol. 2007; 63:125–49. https://doi.org/10.1016/j.critrevonc.2006.12.004.

14. Mangia A, Chiarappa P, Tommasi S, Chiriatti A, Petroni S, Schittulli F, Paradiso A. Genetic heterogeneity by comparative genomic hybridization in BRCAx breast cancers. Cancer Genet Cytogenet. 2008; 182:75–83. https://doi.org/10.1016/j.cancergencyto.2008.01.002.

15. Lacroix M, Leclercq G. The "portrait" of hereditary breast cancer. Breast Cancer Res Treat. 2005; 89:297–304. https://doi.org/10.1007/s10549-004-2172-4.

16. Rahman N, Teare MD, Seal S, Renard H, Mangion J, Cour C, Thompson D, Shugart Y, Eccles D, Devilee P, Meijers H, Nathanson KL, Neuhausen SL, et al. Absence of evidence for a familial breast cancer susceptibility gene at chromosome 8p12-p22. Oncogene. 2000; 19:4170–3. https://doi.org/10.1038/sj.onc.1203735.

17. Imbert A, Chaffanet M, Essioux L, Noguchi T, Adelaide J, Kerangueven F, Le Paslier D, Bonaiti-Pellie C, Sobol H, Birnbaum D, Pebusque MJ. Integrated map of the chromosome 8p12-p21 region, a region involved in human cancers and Werner syndrome. Genomics. 1996; 32:29–38.

18. Nathanson KL, Wooster R, Weber BL. Breast cancer genetics: what we know and what we need. Nat Med. 2001; 7:552–6. https://doi.org/10.1038/87876.

19. Gronwald J, Jauch A, Cybulski C, Schoell B, Bohm-Steuer B, Lener M, Grabowska E, Gorski B, Jakubowska A, Domagala W, Chosia M, Scott RJ, Lubinski J. Comparison of genomic abnormalities between BRCAX and sporadic breast cancers studied by comparative genomic hybridization. Int J Cancer. 2005; 114:230–6. https://doi.org/10.1002/ijc.20723.

20. Jonsson G, Naylor TL, Vallon-Christersson J, Staaf J, Huang J, Ward MR, Greshock JD, Luts L, Olsson H, Rahman N, Stratton M, Ringner M, Borg A, et al. Distinct genomic profiles in hereditary breast tumors identified by array-based comparative genomic hybridization. Cancer Res. 2005; 65:7612–21. https://doi.org/10.1158/0008-5472.CAN-05-0570.

21. Naylor TL, Greshock J, Wang Y, Colligon T, Yu QC, Clemmer V, Zaks TZ, Weber BL. High resolution genomic analysis of sporadic breast cancer using array-based comparative genomic hybridization. Breast Cancer Res. 2005; 7:R1186-98. https://doi.org/10.1186/bcr1356.

22. Nessling M, Richter K, Schwaenen C, Roerig P, Wrobel G, Wessendorf S, Fritz B, Bentz M, Sinn HP, Radlwimmer B, Lichter P. Candidate genes in breast cancer revealed by microarray-based comparative genomic hybridization of archived tissue. Cancer Res. 2005; 65:439–47.

23. Fridlyand J, Snijders AM, Ylstra B, Li H, Olshen A, Segraves R, Dairkee S, Tokuyasu T, Ljung BM, Jain AN, McLennan J, Ziegler J, Chin K, et al. Breast tumor copy number aberration phenotypes and genomic instability. BMC Cancer. 2006; 6:96. https://doi.org/10.1186/1471-2407-6-96.

24. van Beers EH, van Welsem T, Wessels LF, Li Y, Oldenburg RA, Devilee P, Cornelisse CJ, Verhoef S, Hogervorst FB, van't Veer LJ, Nederlof PM. Comparative genomic hybridization profiles in human BRCA1 and BRCA2 breast tumors highlight differential sets of genomic aberrations. Cancer Res. 2005; 65:822–7.

25. Didraga MA, van Beers EH, Joosse SA, Brandwijk KI, Oldenburg RA, Wessels LF, Hogervorst FB, Ligtenberg MJ, Hoogerbrugge N, Verhoef S, Devilee P, Nederlof PM. A non-BRCA1/2 hereditary breast cancer sub-group defined by aCGH profiling of genetically related patients. Breast Cancer Res Treat. 2011; 130:425–36. https://doi.org/10.1007/s10549-011-1357-x.

26. Alvarez C, Aravena A, Tapia T, Rozenblum E, Solis L, Corvalan A, Camus M, Alvarez M, Munroe D, Maass A, Carvallo P. Different Array CGH profiles within hereditary breast cancer tumors associated to BRCA1 expression and overall survival. BMC Cancer. 2016; 16:219. https://doi.org/10.1186/s12885-016-2261-x.

27. Letessier A, Sircoulomb F, Ginestier C, Cervera N, Monville F, Gelsi-Boyer V, Esterni B, Geneix J, Finetti P, Zemmour C, Viens P, Charafe-Jauffret E, Jacquemier J, et al. Frequency, prognostic impact, and subtype association of 8p12, 8q24, 11q13, 12p13, 17q12, and 20q13 amplifications in breast cancers. BMC Cancer. 2006; 6:245. https://doi.org/10.1186/1471-2407-6-245.

28. Yang ZQ, Streicher KL, Ray ME, Abrams J, Ethier SP. Multiple interacting oncogenes on the 8p11-p12 amplicon in human breast cancer. Cancer Res. 2006; 66:11632–43. https://doi.org/10.1158/0008-5472.CAN-06-2946.

29. Bernard-Pierrot I, Gruel N, Stransky N, Vincent-Salomon A, Reyal F, Raynal V, Vallot C, Pierron G, Radvanyi F, Delattre O. Characterization of the recurrent 8p11-12 amplicon identifies PPAPDC1B, a phosphatase protein, as a new therapeutic target in breast cancer. Cancer Res. 2008; 68:7165–75. https://doi.org/10.1158/0008-5472.CAN-08-1360.

30. Gelsi-Boyer V, Orsetti B, Cervera N, Finetti P, Sircoulomb F, Rouge C, Lasorsa L, Letessier A, Ginestier C, Monville F, Esteyries S, Adelaide J, Esterni B, et al. Comprehensive profiling of 8p11-12 amplification in breast cancer. Mol Cancer Res. 2005; 3:655–67. https://doi.org/10.1158/1541-7786.MCR-05-0128.

31. Chen H, Singh RR, Lu X, Huo L, Yao H, Aldape K, Abraham R, Virani S, Mehrotra M, Mishra BM, Bousamra A, Albarracin C, Wu Y, et al. Genome-wide copy number aberrations and HER2 and FGFR1 alterations in primary breast cancer by molecular inversion probe

microarray. Oncotarget. 2017; 8:10845–57. https://doi.org/10.18632/oncotarget.14802.

32. Luo J, Liu S, Leung S, Gru AA, Tao Y, Hoog J, Ho J, Davies SR, Allred DC, Salavaggione AL, Snider J, Mardis ER, Nielsen TO, et al. An mRNA Gene Expression-Based Signature to Identify FGFR1-Amplified Estrogen Receptor-Positive Breast Tumors. J Mol Diagn. 2017; 19:147–61. https://doi.org/10.1016/j.jmoldx.2016.09.007.

33. Shi YJ, Tsang JY, Ni YB, Chan SK, Chan KF, Tse GM. FGFR1 is an adverse outcome indicator for luminal A breast cancers. Oncotarget. 2016; 7:5063–73. https://doi.org/10.18632/oncotarget.6563.

34. Ni HJ, Chang YN, Kao PH, Chai SP, Hsieh YH, Wang DH, Fong JC. Depletion of SUMO ligase hMMS21 impairs G1 to S transition in MCF-7 breast cancer cells. Biochim Biophys Acta. 2012; 1820:1893–900. https://doi.org/10.1016/j.bbagen.2012.08.002.

35. Lundgren K, Holm K, Nordenskjold B, Borg A, Landberg G. Gene products of chromosome 11q and their association with CCND1 gene amplification and tamoxifen resistance in premenopausal breast cancer. Breast Cancer Res. 2008; 10:R81. https://doi.org/10.1186/bcr2150.

36. Sheen-Chen SM, Huang CY, Liu YY, Huang CC, Tang RP. Cortactin in breast cancer: analysis with tissue microarray. Anticancer Res. 2011; 31:293–7.

37. Ni QF, Yu JW, Qian F, Sun NZ, Xiao JJ, Zhu JW. Cortactin promotes colon cancer progression by regulating ERK pathway. Int J Oncol. 2015; 47:1034–42. https://doi.org/10.3892/ijo.2015.3072.

38. Karlsson E, Waltersson MA, Bostner J, Perez-Tenorio G, Olsson B, Hallbeck AL, Stal O. High-resolution genomic analysis of the 11q13 amplicon in breast cancers identifies synergy with 8p12 amplification, involving the mTOR targets S6K2 and 4EBP1. Genes Chromosomes Cancer. 2011; 50:775–87. https://doi.org/10.1002/gcc.20900.

39. Tasian SK, Loh ML. Understanding the biology of CRLF2-overexpressing acute lymphoblastic leukemia. Crit Rev Oncog. 2011; 16:13–24.

40. Jacot W, Fiche M, Zaman K, Wolfer A, Lamy PJ. The HER2 amplicon in breast cancer: Topoisomerase IIA and beyond. Biochim Biophys Acta. 2013; 1836:146–57. https://doi.org/10.1016/j.bbcan.2013.04.004.

41. Kamieniak MM, Munoz-Repeto I, Rico D, Osorio A, Urioste M, Garcia-Donas J, Hernando S, Robles-Diaz L, Ramon YCT, Cazorla A, Saez R, Garcia-Bueno JM, Domingo S, et al. DNA copy number profiling reveals extensive genomic loss in hereditary BRCA1 and BRCA2 ovarian carcinomas. Br J Cancer. 2013; 108:1732–42. https://doi.org/10.1038/bjc.2013.141.

42. Rothe A, Jachimowicz RD, Borchmann S, Madlener M, Kessler J, Reiners KS, Sauer M, Hansen HP, Ullrich RT, Chatterjee S, Borchmann P, Yazaki P, Koslowsky TC, et al. The bispecific immunoligand ULBP2-aCEA redirects natural killer cells to tumor cells and reveals potent anti-tumor activity against colon carcinoma. Int J Cancer. 2014; 134:2829–40. https://doi.org/10.1002/ijc.28609.

43. Ramutton T, Buccheri S, Dieli F, Todaro M, Stassi G, Meraviglia S. γδ T cells as a potential tool in colon cancer immunotherapy. Immunotherapy. 2014; 6:989–99. https://doi.org/10.2217/imt.14.59.

44. Xia Y, Chen B, Shao X, Xiao W, Qian L, Ding Y, Ji M, Gong W. Treatment with a fusion protein of the extracellular domains of NKG2D to IL-15 retards colon cancer growth in mice. J Immunother. 2014; 37:257–66. https://doi.org/10.1097/CJI.0000000000000033.

45. Castells A, Ino Y, Louis DN, Ramesh V, Gusella JF, Rustgi AK. Mapping of a target region of allelic loss to a 0.5-cM interval on chromosome 22q13 in human colorectal cancer. Gastroenterology. 1999; 117:831–7.

46. Cliby W, Ritland S, Hartmann L, Dodson M, Halling KC, Keeney G, Podratz KC, Jenkins RB. Human epithelial ovarian cancer allelotype. Cancer Res. 1993; 53:2393–8.

47. Zweemer RP, Ryan A, Snijders AM, Hermsen MA, Meijer GA, Beller U, Menko FH, Jacobs IJ, Baak JP, Verheijen RH, Kenemans P, van Diest PJ. Comparative genomic hybridization of microdissected familial ovarian carcinoma: two deleted regions on chromosome 15q not previously identified in sporadic ovarian carcinoma. Lab Invest. 2001; 81:1363–70.

48. Chong GO, Jeon HS, Han HS, Son JW, Lee YH, Hong DG, Lee YS, Cho YL. Differential MicroRNA Expression Profiles in Primary and Recurrent Epithelial Ovarian Cancer. Anticancer Res. 2015; 35:2611–7.

49. Palmero EI, Galvao HC, Fernandes GC, Paula AE, Oliveira JC, Souza CP, Andrade CE, Romagnolo LG, Volc S, C Neto M, Sabato C, Grasel R, Mauad E, et al. Oncogenetics service and the Brazilian public health system: the experience of a reference Cancer Hospital. Genet Mol Biol. 2016; 39:168–77. https://doi.org/10.1590/1678-4685-GMB-2014-0364.

50. van Beers EH, Joosse SA, Ligtenberg MJ, Fles R, Hogervorst FB, Verhoef S, Nederlof PM. A multiplex PCR predictor for aCGH success of FFPE samples. Br J Cancer. 2006; 94:333–7. https://doi.org/10.1038/sj.bjc.6602889.

ANEXO III – Arquivo PDF do manuscrito submetido na revista científica "*Cancer*".

Cancer

Whole-exome sequencing of Brazilian non-BRCA1/BRCA2 mutation carrier cases at high-risk for hereditary breast/ovarian cancer

Journal:	*Cancer*
Manuscript ID	Draft
Wiley - Manuscript type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Felicio, Paula; Cancer Hospital of Barretos, Molecular Oncology Research Center Grasel, Rebeca; Cancer Hospital of Barretos, Molecular Oncology Research Center Campacci, Natalia; Cancer Hospital of Barretos, Molecular Oncology Research Center Galvão, Henrique; Cancer Hospital of Barretos, Department of Oncogenetics Torrezan, Giovana; ACCamargo Cancer Center, Genomics and Molecular Biology Group, CIPE de Paula, Andre; Cancer Hospital of Barretos, Center of Molecular Diagnosis Sabato, Cristina; Cancer Hospital of Barretos, Center of Molecular Diagnosis Fernandes, Gabriela; Cancer Hospital of Barretos, Center of Molecular Diagnosis Souza, Cristiano; Cancer Hospital of Barretos, Department of Oncogenetics Michelli, Rodrigo; Cancer Hospital of Barretos, Department of Oncogenetics Andrade, Carlos Eduardo; Cancer Hospital of Barretos, Departament of Oncogenetics Revil, Timothée; McGill University, Department of Human Genetics Ragoussis, Jiannis; McGill University, Department of Human Genetics; McGill University, McGill Genome Centre Couch, Fergus; Mayo Clinic, Department of Laboratory Medicine and Pathology; Mayo Clinic, Department of Health Sciences Research Hart, Steven; Mayo Clinic, Department of Laboratory Medicine and Pathology; Mayo Clinic, Department of Health Sciences Research Reis, Rui; Cancer Hospital of Barretos, Molecular Oncology Research Center; Cancer Hospital of Barretos, Center of Molecular Diagnosis; University of Minho, Life and Health Sciences Research Institute (ICVS), Medical School; ICVS/3B's-PT Government Associate Laboratory Tonin, Patricia; McGill University, Department of Human Genetics; McGill University, Department of Medicine; Research Institute of the McGill University Health Centre, Cancer Research Program Melendez, Matias; Cancer Hospital of Barretos, Molecular Oncology Research Center

	Carraro, Dirce; ACCamargo Cancer Center, Genomics and Molecular Biology Group, CIPE; ACCamargo Cancer Center, Genomic Diagnostic Center Palmero, Edenir; Cancer Hospital of Barretos, Molecular Oncology Research Center; Cancer Hospital of Barretos, Center of Molecular Diagnosis; FACISB, Barretos School of Health Sciences
Keywords:	BRCAX, non-BRCA, whole-exome sequencing, hereditary breast and ovarian cancer predisposition syndrome, familial cancer

Whole-exome sequencing of Brazilian non-*BRCA1/BRCA2* mutation carrier cases at high-risk for hereditary breast/ovarian cancer

Paula Silva Felicio[1]; Rebeca Silveira Grasel[1]; Natalia Campacci[1]; Henrique de Campos Reis Galvão[2]; Giovana Tardin Torrezan[3]; Andre Escremim de Paula[4]; Cristina da Silva Sabato[4]; Gabriela Carvalho Fernandes[4]; Cristiano de Pádua Souza[2]; Rodrigo Depieri Michelli[2]; Carlos Eduardo Andrade[2]; Timothée Revil[5]; Jiannis Ragoussis[5,6]; Fergus J. Couch[7,8]; Steven N. Hart[7,8]; Rui M. Reis[1,4,9,10] ; Patricia N. Tonin[5,11,12]; Matias Eliseo Melendez[1]; Dirce Maria Carraro[3,13]; Edenir Inêz Palmero[1,4,14].

[1]Molecular Oncology Research Center, Barretos Cancer Hospital, Barretos, Brazil.

[2]Department of Oncogenetics, Barretos Cancer Hospital, Barretos, Brazil.

[3]Genomics and Molecular Biology Group, CIPE - A. C. Camargo Cancer Center, São Paulo, Brazil.

[4]Center of Molecular Diagnosis, Barretos Cancer Hospital, Barretos, São Paulo, Brazil.

[5]Department of Human Genetics, McGill University, Montreal, Canada.

[6] McGill Genome Centre, University of McGill, Montreal, Canada.

[7]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota.

[8]Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota.

[9]Life and Health Sciences Research Institute (ICVS), Medical School, University of Minho, Braga, Portugal;

[10]ICVS/3B's-PT Government Associate Laboratory, Braga/Guimarães, Portugal.

[11]Department of Medicine, McGill University, Montreal, Canada.

[12]Cancer Research Program, The Research Institute of the McGill University Health Centre, Montreal, Canada.

[13]Genomic Diagnostic Center, AC Camargo Cancer Center, São Paulo, Brazil.

[14]Barretos School of Health Sciences, Dr. Paulo Prata - FACISB, Barretos, Brazil.

Corresponding author:

Edenir Inêz Palmero, PhD

Molecular Oncology Research Center, Barretos Cancer Hospital

Rua Antenor Duarte Villela, 1331

CEP 14784-400, Barretos, São Paulo, Brazil

Telephone: +551733216600 ext 7057

E-mail: edenirip@yahoo.com.br

Manuscript type: Article.

Tables: 2

Figures: 3

Conflict of Interest

The authors declare no conflict of interest.

Funding

This project is funded through grants from FAPESP (2013/24633-2), PRONON (25000.056766/2015-64), Coordination for the Improvement of Higher Education Personnel (CAPES/PROSUP) and McGill University (Department of Medicine and Cancer Research Program). EIP and RMR are recipients of CNPq Productivity Grants.

Author Contributions

All authors read and approved the final manuscript. Many authors played multiple roles across these activities. Specifically, PSF performed the molecular analysis and interpretation of WES data and written the original draft of the article. RSG and NC recruited patients, collected clinical and family history information and did the review and editing of the manuscript. GTT, TR, JR, FJC and SNH helped in the analysis and interpretation of the data obtained by the WES. HCRG, CPS, RDM, CEA attended and recruited the patients and did the review and editing of the manuscript. AEP, CSS and GCF separated the samples and analyzed *BRCA1*, *BRCA2* and *TP53* genes. RMR, PNT, MEL, DMC contributed to the design of the research and did the review and editing of the manuscript. EIP was responsible for funding and coordination of the project. EIP also performed supervision and final review and edition of the manuscript.

Abstract

The current study aimed to identify, through whole-exome sequencing (WES) of germline/constitutional DNA, new breast and/or ovarian cancer predisposition genes in Brazilian families at high-risk for hereditary breast/ovarian cancer (HBOC). We performed WES in 52 non-*BRCA1/BRCA2* mutation carrier women with high-risk for HBOC. All identified variants were classified using information from population and disease specific databases, as well as *in silico* prediction tools and the ACMG criteria. The majority of patients (n=32, 61.5%) with breast/ovarian cancers were diagnosed at early ages (≤45 years old). Moreover, 87% (n=45) and 31% (n=16) of women reported the presence of a family history of breast or ovarian cancers, respectively. The WES analysis showed the presence of a total of 53 unique loss of function variants and 128 unique rare missense variants. Of these, 23 genes were described as cancer hallmarks by the COSMIC database, including *ATM*, *CHEK2*, *PMS2*, *KRAS*, *KIT* and *EGFR*. Furthermore, variants in the RAD gene family were observed, such as *RAD50*, *RAD51C* and *RAD54L*. In conclusion this is the largest Brazilian WES study involving families at high-risk for HBOC that has brought new insights of the role of potentially new genetic risk factors in Brazilian HBOC.

Keywords: BRCAX, non-BRCA, whole-exome sequencing, hereditary breast and ovarian cancer predisposition syndrome.

INTRODUCTION

Germline variants in *BRCA1/BRCA2* are responsible for approximately 25% of the familial breast cancer (BC) and ovarian cancer (OC) cases, and predispose to the hereditary breast and ovarian cancer (HBOC) syndrome worldwide[1,2]. In Brazil, our group recently reported similar findings, with 21.5% of the 349 index cases with clinical criteria for HBOC syndrome harboring *BRCA1/BRCA2* germline variants[3]. Genomic advances, such as next generation DNA sequencing platforms, allows the analysis of gene panels and the subsequent association of other high and moderated risk genes for HBOC with hereditary BC and OC development. These genes include, among others, *ATM*, *BRIP1*, *CDH1*, *PALB2*, *PTEN*, *RAD51C*, *STK11* and *TP53*[4]. However, for a large proportion of HBOC families (50-80%)[2,5] the genetic cause associated with the BC and OC family history is unknown. A deeper genomic analysis, such as using whole-exome sequencing (WES) followed by bioinformatic analysis of rare variants might reveal new cancer predisposing alleles.

WES can be effective in the diagnosis of individuals for whom the traditional approaches were not conclusive. Besides, it can decreases the exhausting diagnostic odyssey in families with rare or uncharacterized diseases[6]. Additionally, WES can provide opportunities for personalized health-care strategies, such as prevention or early detection of diseases[7]. Several researchers support the idea that WES is the most appropriate tool for identifying genetic familial syndromes. According to Snape *et al.*[8], the analysis of WES offers the potential to perform a strategy not based only on a candidate gene, but on the information available throughout the WES, in search for possible variants with potential significance for the disease ("agnostic approach"). Moreover, as reported by Cooper *et al.*[9], 85% of the mutations relevant to human

are located in the coding region or in canonical splice sites. Therefore, WES could be a promising strategy for identifying new genes associated with an increased risk of BC and OC in high-risk families that do not harbor known pathogenic germline variants in *BRCA1* and *BRCA2*[10].

The aim of the current study was to perform WES in 52 unrelated Brazilian women with high-risk for BC and OC, previously tested negative for pathogenic *BRCA1/BRCA2* germline variants in order to identify driver genes of HBOC.

MATERIALS AND METHODS

Ethical criteria

The present study was approved by Barretos Cancer Hospital (BCH) ethics committee (approval number: 916/2015), and all participants gave their consent to participate.

Patients

Fifty-two unrelated Brazilian women at-risk for HBOC attended at the Oncogenetics Department of BCH[11] were included. All cases had a personal and family history of BC and/or OC. Analysis of the presence of germline variants in *BRCA1/BRCA2/TP53* genes was conducted at the Center of Molecular Diagnosis of BCH as part of routine care through Sanger/NGS sequencing followed by rearrangement analysis through MLPA (Multiplex Ligation-dependent Probe Amplification Analysis), as described elsewhere by Fernandes *et al.* (2016)[3]. Clinical information was obtained through detailed review of the patient's clinical chart. For family history of cancer data, all pedigrees were studied.

DNA isolation and quantification for WES

Genomic DNA was isolated from peripheral blood lymphocytes using the QIAamp DNA Blood Mini Kit (*Qiagen*) following the manufacturer's instructions. DNA concentration was determined using Qubit dsDNA HS Assay Kit (*Thermo Fisher Scientific*).

Whole-exome sequencing

For the exome library preparation, 50ng of constitutional DNA of each sample was used. The kit used was the Nextera Rapid Capture Expanded Exome (*Illumina*),

according to the manufacturer's recommendations. Quantification of the enriched library was performed with Qubit fluorometer (*Thermo Fisher Scientific*) and library size distribution was measured with Agilent Bioanalyzer 2100 (*Agilent Genomics*). Quantified DNA library was loaded on flow cell for subsequent cluster generation. Samples were paired-end sequenced on Illumina NextSeq 500 High Output Kit - 300 cycles (*Illumina*).

Whole-exome sequencing analysis

Briefly, reads were quality trimmed using the Trimmomatic v0.33[12], and then aligned with the genome of reference (UCSC GRCh37/hg19) using the Burrows-Wheeler Aligner (BWA) v0.7.5a. PCR duplicates were removed using Picard v1.106 and BAM files were processed using the Genome Analysis Toolkit (GATK) v2.7.2 software. Realignment and search for indels were performed using GATK HaplotypeCaller and annotated using snpEFF v4.3 and SnpSift[13]. A GEMINI v.0.19.1 database was created[14], and variants selected per functional rules. Additionally, variants described by snpEFF/GEMINI as "low-impact" were removed since they are assumed to have benign effects on DNA or protein behavior.

The analysis workflow is illustrated in Figure 1. For a function-based prioritization, variants leading to loss of function ("high-impact" variants: frameshift, nonsense, and canonical splice site variants) and missense variants (classified as "medium-impact" variants) were selected. For quality filtering, variants with vertical coverage ≥10x and variant allele frequency (VAF) ≥0.25 were selected. Next, a total of 2,319 cancer-associated genes were analyzed (described below). Variants present in the population database Genome Aggregation Database (gnomAD[15]), with a frequency ≤2% (minor

allele frequency [MAF] ≤0.02) were maintained. Furthermore, a recently publicly available Brazilian database of WES from 609 healthy individuals (AbraOM—Brazilian genomic variants[16]) was also used for manually excluding population-specific variants (MAF ≤0.02 were maintained). Loss of function variants were manually examined with Integrative Genomics Viewer (IGV)[17] to remove possible artifacts.

Cancer gene reference lists

Three databases were used to generate a candidate list of genes previously reported associated with any type of cancer, namely: *i)* The Cancer Gene Census v.86, a set of 719 genes manually curated by the Sanger Institute[18], *ii)* a query of DISEASES[19], a database of disease-gene associations based largely on text-mining approaches, and, *iii)* UniprotKB[20], a manually curated database of protein functions (using the keyword-terms "cancer", "tumor-suppressor gene", "proto-oncogene" and "oncogene"). From these databases, a reference list of 2,319 genes was generated for prioritizing and characterizing gene variants. Detailed information about these genes is available in Table S1.

RESULTS

1. Clinico-pathological characterization

Of the 52 Brazilian women analyzed, 41 had BC, nine had OC and two were initially diagnosed with OC which was followed by a BC diagnosis. Detailed clinico-pathological features are specified in Table S2. Briefly, the average age of cancer diagnosis was 41.1 years, ranging from 20 to 60 years. The majority of BC cases had invasive ductal carcinoma (85.4%), and were estrogen and progesterone positive (70.3% and 60.5%, respectively) and HER2 negative (77.8%) tumors. Most of the BC cases were molecularly classified as luminal type tumors (73.0%), seven were triple negative tumors (18.9%) and three (8.1%) were tumors overexpressing HER2. The majority of women with OC developed high-grade serous adenocarcinoma subtype tumors (90.9%).

2. Family history

The family history of the patients included in this study is depicted in Table S3. Besides BC and OC, other cancers associated with HBOC spectrum, such as cancers of the prostate (n=14 relatives, 10 families) and pancreas (n=4 relatives, 4 families) were observed. BC cases were observed in first degree relatives in 56% of the families (n=29). In addition, the presence of OC among first degree relatives was observed in four families, and in three of these four families the proband was also diagnosed with OC.

3. Germline variants by whole-exome sequencing

The WES identified a total of 2,536,915 variants in the 52 cases (Figure 1). To narrow down the analysis, a reference list of 2,319 candidate genes previously associated with cancer was used. After applying the depicted workflow (Figure 1), a total of 3,027 variants were selected for further bioinformatic analyses. Next, variants found in ≥25% of our cohort (13 patients) were excluded (totaling 183 variants) as they likely represent sequencing and/or annotation artefacts. Therefore, a total of 2,844 variants were analyzed in this study. Of these, 75 (2.7%) were classified as loss of function (LoF) and 2,769 (97.3%) were missense/inframe variants.

The frequency of variants in each case is detailed in Figure 2. Overall, an average of 55 variants was observed, ranging from 15 to 131 variants per patient. In addition, when the variants were grouped by the impact, an average of 1 LoF variant and 53 missense/inframe variants were identified in each patient. Moreover, germline variants described by ClinVar as benign/likely-benign (B/LB) and pathogenic/likely-pathogenic (P/LP) were analyzed separately. Of these, 46 variants were described as B/LB by ClinVar and, only one variant with LoF classified as LP/P (without conflicting interpretations) by ClinVar. The frameshift variant classified as pathogenic by ClinVar (*PMS2*: c.2182_2184delinsG) was identified in patient 306, who had melanoma at 26 years of age and BC at 36 years of age.

3.1. Germline variants with conflicting interpretations, VUS or not classified by ClinVar

Variants with discordance and/or not reported in the ClinVar database were evaluated separately. A total of 2,797 variants were identified in this part of the analysis, including 68 LoF and 2,729 missense/inframe variants. An average of 52 variants in

each case was observed, ranging from 13 to 131 variants. The frequency and the type of variants found in each case, after applying the filters are detailed in Figure 2.

3.1.1. LoF variants

After applying all filters described above, a total of 52 unique LoF variants (with conflicting interpretations, VUS or not classified by ClinVar) were identified (19 frameshift, 23 nonsense and 10 splice acceptor/donor variant). All LoF variants identified in genes associated with cancer are detailed in Table 1 and Figure S1.

The most frequently mutated gene with LoF variants was *CCND3*. A unique nonsense variant identified in *CCND3* gene (c.379G>T) was observed in four unrelated women (patients ID: 80, 85, 179 and 1014) with BC (average age at diagnosis: 46 years). All of *CCND3* variant carriers reported a family history of BC. The presence of this mutation in the four families from each variant carrier was confirmed by Sanger sequencing.

Variants were identified in in genes described by the COSMIC database as "hallmarks of cancer", such as *AFF3*, *DROSHA* and *SLC34A2* (Figure 3). The frameshift variant identified in the *AFF3* (c.1788del) has not been described by ClinVar. The patient (ID 1097) identified with this variant had OC at 46 years of age and reported a family history of BC. The nonsense variant in *DROSHA* (c.1498G>T), also not reported by ClinVar, was identified in a patient (ID 133) diagnosed with BC at 46 years of age who also reported a family history of BC. The splice acceptor variant in *SLC34A2* (c.113-2A>G), not described by ClinVar, was identified in a patient (ID 565) who reported a personal/family history of OC.

Besides, identifying variants in genes known associated with other types of hereditary/familial tumors, LoF variants in other cancer associated genes such as

FANCA, *FLCN*, *POLQ*, *PTCH1*, *RAD51C*, *RAD54B* and *TSC2* were also identified in our high-risk cases. The variant identified in *POLQ*, (c.4262_4268del), was found in two unrelated BC patients. One of them (patient ID 689) had bilateral BC at 47 years of age and reported a family history of leukemia, BC and colorectal tumors. The other patient (ID 638) with the same *POLQ* variant, had BC at 42 years of age and described a family history BC and gastric tumors. The frameshift variant identified in *RAD51C* (c.896_907del) was observed in patient 656, who had OC at 41 years of age and BC at 48 years of age, and reported a second degree relative with OC. A frameshift variant in another gene in the RAD-family, the *RAD54B* (c.2733del), was identified in a patient (ID 19) with BC at 27 years of age, who reported two relatives with BC (before 45 years of age) in their family history. The frameshift variant identified in the *TSC2* (c.5068+2_5068+6dupTAGGG) was observed in a patient (1186) who had OC at 20 years of age and, one year later, was diagnosed with an appendiceal neuroendocrine tumor. This patient reported a second-degree relative uterine cancer. The nonsense variants identified in *FLCN* (c.1029A>C) was observed in patient 65, who had a personal/familial of BC. The variant identified in *PTCH1* (c.1A>C) was detected in patient 1055, who was diagnosed with OC at 57 years of age and reported the presence of BC in their family history. The splice donor variant involved in the *FANCA* gene (c.435+2dup) was observed in patient 80. This woman had BC at 43 years of age and reported six relatives with BC in her family history. The significance of any of above described LoF variants are unknown as none of them have been reported in the ClinVar database.

3.1.2. Missense variants

To further evaluate the missense variants, three bioinformatic tools were selected and applied in the analysis of our variants, namely CADD (score ≥20), REVEL (score ≥0.5) and M-CAP (score ≥0.025). The number of variants de-prioritized by each tool is detailed in Figure S2. As a result of this filtering step, a total of 136 rare missense variants were selected (128 unique variants) for further evaluation. There was an average of three variants per patient, ranging from 1 to 8 variants. All selected missense variants that were considered pathogenic by the three bioinformatic tools, together with information regarding the patient´s age and cancer type at diagnosis are detailed in Table 2 and Figure S1.

The genes most frequently altered were *MPP3* and *PMS1*. A variant in *MPP3* (c.617C>A), not described by ClinVar, was identified in three unrelated patients (IDs: 974, 1096 and 1231), where two had BC at 46 years of age and one had OC at 21 years of age. All patients with this *MPP3* variant reported a family history of BC. Two unique variants were identified in *PMS1* (Table 2), both found in patients who reported a family history of OC.

We identified a total of 21 unique missense variants in genes described as "hallmarks of cancer" by COSMIC database (Figure 3). Among these genes, we can highlight *CHEK2*, *FAT1*, *ATM*, *EGFR*, *ERBB4*, *KIT*, *KRAS* and *TERT*. Some of these genes have already been associated with hereditary/familial tumors, such as the *ATM* and *CHEK2* genes. The variants identified in *CHEK2* have conflicting interpretations of pathogenicity at ClinVar. The missense *CHEK2* variant (c.349A>G; ClinVar: pathogenic/ likely-pathogenic) was identified in a patient with BC at 35 years of age (ID= 1326) and a family history of three BC cases. The other variant in *CHEK2* (c.538C>T; ClinVar: VUS/ likely-benign/ benign) was identified in a patient with OC at 21 years of age (ID= 695),

who reported a family history of BC and colorectal cancer. Two different variants were identified in *FAT1* (IDs: 494: c.7957G>A and 960: c.9583T>A), where neither have been described by ClinVar. The variant identified in *ATM* (c.1595G>A) having conflicting interpretations of pathogenicity by ClinVar, was identified in a patient with BC at 37 years of age, and, thyroid cancer at 39 years of age (ID=1327). In addition, this patient reported two relatives BC and one with leukemia. The variant identified in the oncogene *EGFR* (c.352G>T), which was not reported by ClinVar, was identified in patient 1231 with OC at 21 years of age who reported ne case of BC in a first-degree relative. In patient 29, variants in the oncogenes *ERBB4* (c.3446G>T) and *KRAS* (c.461A>G) were observed, where neither have been reported by ClinVar. This patient had OC at 42 years of age and BC at 53 years of age. In addition, this patient reported the presence of several tumors in her family history, especially in first- and second-degree relatives, notably BC, OC, uterine and gastric cancers. The variant found in the oncogene *KIT* (c.50T>C), reported by ClinVar as VUS, was observed in patient 426. This patient had BC at 38 years of age, and reported a family history of BC and colorectal tumors. The variant identified in *TERT* (c.3332C>T), reported by ClinVar as VUS, was identified in patient 275. This patient had OC at 60 years of age and reported a family history of BC and prostate cancer.

Regarding the genes associated with DNA repair, variants in *EXO1*, *FAN1*, *MLH1*, *MC1R*, *RAD54L* and *RAD50* were identified in our cohort. Two patients (ID: 638 and 1095), both with BC before the age of 45 shared the same variant in *EXO1* (c.820G>A). Two unrelated patients (426 and 1264) carried the same *FAN1* variant (c.149T>G), and this variant has not been described by ClinVar. Additionally, these patients had BC at 38 and 48 years of age, respectively. The variant identified in *MLH1* (c.794G>A), reported

by ClinVar as VUS, was observed in patient 85, who had BC at 51 years of age and reported a family history of BC and gastric cancer. The same *MC1R* variant (c.4604G>A) was observed in patients 960 and 1014. This variant was described by ClinVar as having uncertain significance and was associated with cardiovascular disease. Both *MC1R* variant carriers had BC (patient 960 had two primary BC) and reported a family history of BC. Interestingly, two different missense variants in *RAD54L*, neither reported by Clinvar, were identified in two OC patients (patient 320: c.604C>T and patient 565: c.1094G>A). Both carriers reported a family history of OC. Although only one variant was observed in *RAD50*, a gene involved in DNA repair, this result is noteworthy. The variant (c.353T>C) observed in this study in a patient (1097) with OC at 46 years of age, was reported by ClinVar as uncertain significance, with only two submitters. Details about the personal and family history of the patients with variants in these genes, as well as details regarding the type of variant identified are depicted in Table 2.

Discussion

For a great proportion of the non-*BRCA1/BRCA2* patients with familial BC and/or OC the genetic cause associated with their cancer predisposition is unknown. In the current study, we aimed to identify rare variants that would contribute to HBOC susceptibility in 52 Brazilian high-risk HBOC families *BRCA1/BRCA2* mutation-negative by WES and bioinformatic analyses approach.

An initial investigation in our cohort was based on the protein function, prioritizing the analysis of cancer-associated genes. In addition, the variant classification by ClinVar was considered. Variants classified as benign/likely-benign and pathogenic by ClinVar were removed because there is evidence to support their reported classification

through *in vitro* assay, disease mechanism, population studies or segregation studies[21].

A total of 53 unique LoF and 128 unique rare missense variants were identified in the present study. Germline variants were identified in 23 genes described as "hallmarks of cancer", by COSMIC. Of these, four genes were affected with LoF variants, and, 19 genes were affected by rare missense variants. We emphasize some of these genes due to their genetic function and the possible association with hereditary/familial tumors.

The most frequently mutated gene with LoF variants was *CCND3*, identified in four unrelated BC patients (7.7% of the cohort). *CCND3* encodes a protein that functions in the regulation of cyclin-dependent kinases in the cell cycle[20]. An independent study, reported the presence of the same variant (c.379G>T), in germline tissue, in 1.61% (9/557) of the Caucasians OC patients[22]. Germline variants of *CCND3* have not been reported in BC, though 15% (n= 3/20) of metastatic metaplastic breast carcinoma, had an amplification involving *CCND3*.Interestingly, the authors of that study suggest that *CCND3* could be a potential target of therapy (Nutlins)[23].

The nonsense variant (c.1498G>T) identified in *DROSHA*, was not reported by ClinVar or in the literature. *DROSHA* has been described as crucial in microRNA biogenesis and, more recently, in translational control and in the direct interaction with p53 effectors associated with RNA binding[20]. Additionally, some authors have described its protein interaction with BRCA1 and involvement in the *SMAD3/TP53/DHX9* pathway, which promotes miRNA maturation[24,25]. Interestingly, a single nucleotide polymorphism in the *DROSHA* gene (rs78393591), was reported in women of African ancestry who had BC[24].Moreover, somatic mutation in *DROSHA* has been shown to be high frequent in

Wilms tumor[26]. These observations suggest a possible role for *DROSHA* in cancer etiology.

A previously not unreported splice acceptor variant identified in the hallmark of cancer-associated gene *SLC34A2* was found in a woman with personal and family history of OC. Study performed by Kanchi *et al.* (2014), through the analysis of germline and somatic variants in OC, observed the presence of a germline splice-site variant (c.1458+2T>C) and a missense variant (c.1079C>T)[22] involved in this suppressor tumor gene. However, as these variants were found at a low in lower frequency in OC cases, the authors could not associate the variant with disease development.

Germline variants in the mismatch DNA repair (MMR) system that involve *MLH1, MSH2, PMS2*, and *MSH6*, have been associated with Lynch syndrome which feature largely feature colorectal and endometrial cancers, and OC[27,28]. However, some researchers have also identified an association of germline pathogenic variants in MMR genes with BC, although a discrepancy is observed in the literature. A study conducted by Couch *et al.* (2017), analyzing 65,057 non-*BRCA1/BRCA2* mutation carrier women-with BC and referred for hereditary cancer genetic testing, estimated moderate risks of BC with carrier mutations in *MLH1* (OR: 1.15, 95% CI: 0.30-4.19), *MSH2* (OR: 2.46, 95% CI: 0.81-6.93), *MSH6* (OR: 1.93, 95% CI: 1.16-3.27) or *PMS2* (OR: 0.82,95% CI: 0.44-1.47)[29]. In the current study, variants involved in genes *MLH1, MSH2, MSH6* (missense variants) and *PMS2* (frameshift) were also observed in patients with personal and family history of BC. Although the majority of the variants identified in these genes in our study were described by ClinVar as VUS, some were considered

pathogenic by submitters this database. Further study is warranted to determine the potential pathogenicity and penetrance of VUS identified in MMR genes for BC risk.

In this study we identified another gene with LoF and previously associated with colorectal cancer in *POLQ* (c.4262_4268del). Interestingly, this was found in two unrelated patients with BC. Studies published by Wang *et al.* (2008)[30] and Brandalize *et al.* (2014)[31] reported the presence of germline variants in *POLQ*, involved patients non-*BRCA1/BRCA2* mutated BC cases with a family history of BC. In addition, a case-control study published by Family *et al.* (2015)[32], associated three missense variants in *POLQ* with an increased risk in developing BC. Thus, the authors concluded that *POLQ* could be considered a possible gene candidate involved in the development of hereditary BC.

Other genes related to hallmarks of cancer have also been associated with an increased risk for the development of BC and OC. *ATM* and *CHEK2* are good examples. The rare missense variants identified in these genes in our cohort were described with conflicting interpretations by ClinVar. In 2017, Couch *et al.* showed that pathogenic variants in *ATM* and *CHEK2* were associated with moderate risk of BC (OR: 2.78, 95%CI: 2.22-3.62 and 2.26, 95%CI: 1.89-2.72, respectively)[29]. Similar results were published by Lu *et al.* (2018). The risk to develop BC carriers of mutations in *ATM* was 2.97 (95%CI: 1.67-5.68) and in *CHEK2* was 2.19 (95%CI: 1.40-3.56). Although the authors of that study also estimated the risks of OC, only *ATM* mutation carriers showed an increased risk (OR: 2.85 - 95%CI: 1.30-6.32) for this cancer [33]. An independent study reported that the risk of OC in *CHEK2* mutation carriers was also moderate (OR: 0.86, 95% CI: 0.56-1.33)[34].

We identified variants in *FAT1*, a tumor suppressor gene involved in the WNT/β-catenin pathway[20], in two families with either BC or OC cases. Variants in *FAT1* was also identified in the germline in WES study of 147 cases of colorectal cancer (CRC) from Canadian and Australian families[35]. All patients in that study were diagnosed under the age of 50 years, and, showed the criteria of Familial Colorectal Cancer Type X (FCCTX)[35]. However, none of our patients with *FAT1* variants reported the presence of colorectal cancer in their family history.

It is important to emphasize that some rare missense variants identified in our study were found in oncogenes, such as *EGFR*, *ERBB4*, *KIT* and *KRAS*, though many of these variants have not been described by ClinVar. These genes are described as kinase superfamily[20]. Somatic amplification and variants in these genes have been molecularly characterized and are established drivers in many cancer types, especially in BC, lung and gastrointestinal stromal cancers, and glioblastomas. However, some reports have described the presence of germline variants in these genes in familial BC/OC. Additionally, Penkert *et al.* (2018), performed next-generation sequencing of 94 cancer predisposition genes in German cohort of 83 unrelated women diagnosed with BC, without germline variants in *BRCA1, BRCA2* and *TP53*, and reported germline missense variants in *EGFR* (c.2039G>A) and *KIT* (c.391G>A), both with frequency of 1.2%[36]. In our cohort, variants in these genes were found in patients with OC (*EGFR*), both OC and BC cancer (*ERBB4* and *KRAS*) and BC (*KIT*). The variants identified in *KRAS* and *EGFR* were located in domains predicted to activate these proto-oncogenes (Ras and Receptor-L domain, respectively)[37]. The association of the variants identified in these oncogenes in our cohort will be further evaluated by co-segregation analysis.

Some variants involving the RAD-family were observed in the current study, were mainly found in probands with personal and family history of OC. The RAD-family is involved in the homologous recombination and DNA repair processes[20].Variants in this gene family included *RAD50* (missense), *RAD54L* (missense), *RAD51C* (frameshift), and *RAD54B* (frameshift) where most have not been reported by ClinVar. Some authors have studied the presence of germline variants in these genes in non-*BRCA1/BRCA2* mutation carrier patients at-high risk for BC/OC. Kanchi *et al.*, (2014) described the presence of germline pathogenic variants involving *RAD51C* and *RAD54L* in OC cases (both with a frequency of 0.23%)[22]. A targeted mutation screen of *RAD51B*, *RAD51C* and *RAD51D* in 3,429 patients with invasive epithelial OC and 2,000 unaffected *BRCA1/BRCA2* mutation negative women, found rare variants in OC cases in *RAD51C* (14, 0.41%), *RAD51D* (12, 0.35%) and in *RAD51B* (two, 0.06%)[29]. In addition, *RAD51C* variants were associated with a risk of 5.2 (OR, 95% CI: 1.1-24.0)[38]. An increased risk for OC was purported for *RAD51C* variant carriers (OR, not estimated) in a study by Lu *et al.* (2018)[33]. In contrast Couch *et al.*, (2017), showed that potentially pathogenic variants in *RAD51C* were not associated with increased risk for BC (OR: 0.78, 95% CI: 0.47-1.37)[29]. A recent study by Fan *et al.* (2018) investigating the clinical impact of *RAD50* germline variants in 7,657 unselected BC non-*BRCA1/BRCA2* carrier patients reported a carrier frequency of 26/7,657 (0.34%) in patients. Interestingly, the same study reported a multivariate analysis revealing that potentially pathogenic *RAD50* variants were an independent unfavorable predictor of recurrence-free survival (OR: 2.66; 95%CI: 1.18-5.98) and disease-specific survival (OR: 4.36; 95%CI: 1.58-12.03)[39]. Thus, the impacts of the variants in *RAD50* and *RAD51C* on BC susceptibility are still controversial. In our study, the *RAD54B* variant was

identified only in a BC case, and the carrier of a variant *RAD51C* developed two primary tumors of the ovary and breast, though this might reflect the rarity of germline carriers of these genes and sample size of cases investigated by WES in our study.

Although WES has become an effective approach to identify variants in alleles and thereby new candidate genetic predisposition genes in many hereditary diseases, it comes with numerous challenges. The different lists of genes that resulted from different breast/ovarian tumors WES studies may be elucidated in part by the different pipelines and bioinformatics tools used to evaluate these data, as well as can be a consequence of the population-specific variability or even due to the selection criteria adopted. Cohorts used to for WES study also play a role in comparing results from different studies. Most notably, we have compared our findings with the Kanchi *et al.* study[22] who reported germline (and somatic) variants in unselected OC cases in contrast to our study of high-risk OC cases from cancer families. Furthermore, WES data users apply different filters to help prioritize variants, thus some variants can be included/removed erroneously.

CONCLUSION

In summary, the present study performed a characterization of germline variants identified in cancer-associated genes, using WES and bioinformatic analyses in Brazilian non-*BRCA1/BRCA2/TP53* mutation-carrier women with BC and/or OC. Our findings suggest that several novel cancer-associated genes also may have a role in HBOC. In addition, the present study provides additional evidence for the association of moderate-risk genes, such as *CHEK2*, *RAD50, RAD51C*, and *RAD54L*, to the development of familial BC/OC. The candidate genes identified in our high-risk cancer

cases would require further validation in Brazilian and other cancer cohorts. Specific variants would also require *in vitro* analysis to investigate functional consequences on protein function. Such advances will help with the molecular cataloguing of breast/ovarian tumors in non-*BRCA1/BRCA2* mutation carrier patients as well as the development of gene-based early detection strategies and targeted therapies as has been established in *BRCA1/BRCA2* mutation carrier women.

Acknowledgements

We would like to thank all our patients and their families who contributed to this study. The authors are very grateful to Suzanna L. Arcand, Wejdan Alenezi and Bruna Durães de Figueiredo Barros for whole-exome sequencing analysis.

Reference

1. Lu, K.H., Wood, M.E., Daniels, M., Burke, C., Ford, J., Kauff, N.D., Kohlmann, W., Lindor, N.M., Mulvey, T.M., Robinson, L., Rubinstein, W.S., Stoffel, E.M., Snyder, C., Syngal, S., Merrill, J.K., Wollins, D.S., and Hughes, K.S. (2014) American Society of Clinical Oncology Expert Statement: Collection and Use of a Cancer Family History for Oncology Providers. *Journal of Clinical Oncology*, **32** (8), 833–840.
2. Couch, F.J., Nathanson, K.L., and Offit, K. (2014) Two decades after BRCA: setting paradigms in personalized cancer care and prevention. *Science*, **343**, 1466–70.
3. Fernandes, G.C., Michelli, R.A., Galvao, H.C., Paula, A.E., Pereira, R., Andrade, C.E., Felicio, P.S., Souza, C.P., Mendes, D.R., Volc, S., Berardinelli, G.N., Grasel, R.S., Sabato, C.S., Viana, D.V., Mauad, E.C., Scapulatempo-Neto, C., Arun, B., Reis, R.M., and Palmero, E.I. (2016) Prevalence of BRCA1/BRCA2 mutations in a Brazilian population sample at-risk for hereditary breast cancer and characterization of its genetic ancestry. *Oncotarget*, **7**, 80465–80481.
4. Lu, C., Xie, M., Wendl, M.C., Wang, J., McLellan, M.D., Leiserson, M.D.M., Huang, K., Wyczalkowski, M.A., Jayasinghe, R., Banerjee, T., Ning, J., Tripathi, P., Zhang, Q., Niu, B., Ye, K., Schmidt, H.K., Fulton, R.S., McMichael, J.F., Batra, P., Kandoth, C., Bharadwaj, M., Koboldt, D.C., Miller, C.A., Kanchi, K.L., Eldred, J.M., Larson, D.E., Welch, J.S., You, M., Ozenberger, B.A., Govindan, R., Walter, M.J., Ellis, M.J., Mardis, E.R., Graubert, T.A., Dipersio, J.F., Ley, T.J., Wilson, R.K., Goodfellow, P.J., Raphael, B.J., Chen, F., Johnson, K.J., Parvin, J.D., and Ding, L. (2015) Patterns and functional implications of rare germline variants across 12 cancer types. *Nature Communications*, **6** (1).

5. Jones, M.R., Kamara, D., Karlan, B.Y., Pharoah, P.D.P., and Gayther, S.A. (2017) Genetic epidemiology of ovarian cancer and prospects for polygenic risk prediction. *Gynecologic Oncology*, **147** (3), 705–713.
6. Berberich, A.J., Ho, R., and Hegele, R.A. (2018) Whole genome sequencing in the clinic: empowerment or too much information? *Canadian Medical Association Journal*, **190** (5), E124–E125.
7. Schwarze, K., Buchanan, J., Taylor, J.C., and Wordsworth, S. (2018) Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the literature. *GENETICS in MEDICINE*.
8. Snape, K., Ruark, E., Tarpey, P., Renwick, A., Turnbull, C., Seal, S., Murray, A., Hanks, S., Douglas, J., Stratton, M.R., and Rahman, N. (2012) Predisposition gene identification in common cancers by exome sequencing: insights from familial breast cancer. *Breast Cancer Research and Treatment*, **134**, 429–433.
9. Cooper DN, Krawczak M, Antonorakis SE (1995) The nature and mechanisms of human gene mutation., in *The Metabolic and Molecular Bases of Inherited Disease.*, Scriver C, Beaudet al., Sly WS, Valle D, editors., New York, pp. 259–291.
10. Gracia-Aznarez, F.J., Fernandez, V., Pita, G., Peterlongo, P., Dominguez, O., de la Hoya, M., Duran, M., Osorio, A., Moreno, L., Gonzalez-Neira, A., Rosa-Rosa, J.M., Sinilnikova, O., Mazoyer, S., Hopper, J., Lazaro, C., Southey, M., Odefrey, F., Manoukian, S., Catucci, I., Caldes, T., Lynch, H.T., Hilbers, F.S.M., van Asperen, C.J., Vasen, H.F.A., Goldgar, D., Radice, P., Devilee, P., and Benitez, J. (2013) Whole Exome Sequencing Suggests Much of Non-BRCA1/BRCA2 Familial Breast Cancer Is Due to Moderate and Low Penetrance Susceptibility Alleles. *Plos One*, **8**.

11. Palmero, E.I., Galvão, H.C.R., Fernandes, G.C., Paula, A.E. de, Oliveira, J.C., Souza, C.P., Andrade, C.E., Romagnolo, L.G.C., Volc, S., C. Neto, M., Sabato, C., Grasel, R., Mauad, E., Reis, R.M., and Michelli, R.A.D. (2016) Oncogenetics service and the Brazilian public health system: the experience of a reference Cancer Hospital. *Genetics and Molecular Biology*, **39** (2), 168–177.

12. Bolger, A.M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30** (15), 2114–2120.

13. Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w ^{1118} ; iso-2; iso-3. *Fly*, **6** (2), 80–92.

14. Paila, U., Chapman, B.A., Kirchner, R., and Quinlan, A.R. (2013) GEMINI: Integrative Exploration of Genetic Variation and Genome Annotations. *PLoS Computational Biology*, **9** (7), e1003153.

15. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., Tukiainen, T., Birnbaum, D.P., Kosmicki, J.A., Duncan, L.E., Estrada, K., Zhao, F., Zou, J., Pierce-Hoffman, E., Berghout, J., Cooper, D.N., Deflaux, N., DePristo, M., Do, R., Flannick, J., Fromer, M., Gauthier, L., Goldstein, J., Gupta, N., Howrigan, D., Kiezun, A., Kurki, M.I., Moonshine, A.L., Natarajan, P., Orozco, L., Peloso, G.M., Poplin, R., Rivas, M.A., Ruano-Rubio, V., Rose, S.A., Ruderfer, D.M., Shakir, K., Stenson, P.D., Stevens, C., Thomas, B.P., Tiao, G., Tusie-Luna, M.T., Weisburd, B., Won, H.-H., Yu, D., Altshuler, D.M., Ardissino, D., Boehnke, M., Danesh, J., Donnelly, S., Elosua, R., Florez, J.C., Gabriel, S.B., Getz, G., Glatt, S.J., Hultman, C.M., Kathiresan, S., Laakso,

M., McCarroll, S., McCarthy, M.I., McGovern, D., McPherson, R., Neale, B.M., Palotie, A., Purcell, S.M., Saleheen, D., Scharf, J.M., Sklar, P., Sullivan, P.F., Tuomilehto, J., Tsuang, M.T., Watkins, H.C., Wilson, J.G., Daly, M.J., MacArthur, D.G., and Exome Aggregation Consortium (2016) Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, **536**, 285.

16. Naslavsky, M.S., Yamamoto, G.L., de Almeida, T.F., Ezquina, S.A.M., Sunaga, D.Y., Pho, N., Bozoklian, D., Sandberg, T.O.M., Brito, L.A., Lazar, M., Bernardo, D.V., Amaro, E., Duarte, Y.A.O., Lebrão, M.L., Passos-Bueno, M.R., and Zatz, M. (2017) Exomic variants of an elderly cohort of Brazilians in the ABraOM database: NASLAVSKY ET AL. *Human Mutation*, **38** (7), 751–763.
17. Thorvaldsdottir, H., Robinson, J.T., and Mesirov, J.P. (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, **14** (2), 178–192.
18. Forbes, S.A., Tang, G., Bindal, N., Bamford, S., Dawson, E., Cole, C., Kok, C.Y., Jia, M., Ewing, R., Menzies, A., Teague, J.W., Stratton, M.R., and Futreal, P.A. (2010) COSMIC (the Catalogue of Somatic Mutations in Cancer): a resource to investigate acquired mutations in human cancer. *Nucleic Acids Res.*, **38** (Database issue), D652-657.
19. Pletscher-Frankild, S., Pallejà, A., Tsafou, K., Binder, J.X., and Jensen, L.J. (2015) DISEASES: Text mining and data integration of disease–gene associations. *Methods*, **74**, 83–89.
20. The UniProt Consortium (2017) UniProt: the universal protein knowledgebase. *Nucleic Acids Res.*, **45** (D1), D158–D169.

21. Bean, L.J.H., and Hegde, M.R. (2017) Clinical implications and considerations for evaluation of in silico algorithms for use with ACMG/AMP clinical variant interpretation guidelines. *Genome Medicine*, **9** (1).

22. Kanchi, K.L., Johnson, K.J., Lu, C., McLellan, M.D., Leiserson, M.D.M., Wendl, M.C., Zhang, Q., Koboldt, D.C., Xie, M., Kandoth, C., McMichael, J.F., Wyczalkowski, M.A., Larson, D.E., Schmidt, H.K., Miller, C.A., Fulton, R.S., Spellman, P.T., Mardis, E.R., Druley, T.E., Graubert, T.A., Goodfellow, P.J., Raphael, B.J., Wilson, R.K., and Ding, L. (2014) Integrated analysis of germline and somatic variants in ovarian cancer. *Nat Commun*, **5**, 3156.

23. Ross, J.S., Badve, S., Wang, K., Sheehan, C.E., Boguniewicz, A.B., Otto, G.A., Yelensky, R., Lipson, D., Ali, S., Morosini, D., Chliemlecki, J., Elvin, J.A., Miller, V.A., and Stephens, P.J. (2015) Genomic profiling of advanced-stage, metaplastic breast carcinoma by next-generation sequencing reveals frequent, targetable genomic abnormalities and potential new treatment options. *Arch. Pathol. Lab. Med.*, **139** (5), 642–649.

24. Qian, F., Feng, Y., Zheng, Y., Ogundiran, T.O., Ojengbede, O., Zheng, W., Blot, W., Ambrosone, C.B., John, E.M., Bernstein, L., Hu, J.J., Ziegler, R.G., Nyante, S., Bandera, E.V., Ingles, S.A., Press, M.F., Nathanson, K.L., Hennis, A., Nemesure, B., Ambs, S., Kolonel, L.N., Olopade, O.I., Haiman, C.A., and Huo, D. (2016) Genetic variants in microRNA and microRNA biogenesis pathway genes and breast cancer risk among women of African ancestry. *Hum. Genet.*, **135** (10), 1145–1159.

25. Kawai, S., and Amano, A. (2012) BRCA1 regulates microRNA biogenesis via the DROSHA microprocessor complex. *J. Cell Biol.*, **197** (2), 201–208.

26. Torrezan, G.T., Ferreira, E.N., Nakahata, A.M., Barros, B.D.F., Castro, M.T.M., Correa, B.R., Krepischi, A.C.V., Olivieri, E.H.R., Cunha, I.W., Tabori, U., Grundy, P.E., Costa, C.M.L., de Camargo, B., Galante, P.A.F., and Carraro, D.M. (2014) Recurrent somatic mutation in DROSHA induces microRNA profile changes in Wilms tumour. *Nat Commun*, **5**, 4039.

27. Lynch, H.T., Snyder, C.L., Shaw, T.G., Heinen, C.D., and Hitchins, M.P. (2015) Milestones of Lynch syndrome: 1895–2015. *Nature Reviews Cancer*, **15**, 181.

28. Buckley, A.R., Ideker, T., Carter, H., Harismendy, O., and Schork, N.J. (2018) Exome-wide analysis of bi-allelic alterations identifies a Lynch phenotype in The Cancer Genome Atlas. *Genome Med*, **10** (1), 69.

29. Couch, F.J., Shimelis, H., Hu, C., and et al (2017) Associations between cancer predisposition testing panel genes and breast cancer. *JAMA oncology*.

30. Wang, X., Szabo, C., Qian, C., Amadio, P.G., Thibodeau, S.N., Cerhan, J.R., Petersen, G.M., Liu, W., and Couch, F.J. (2008) Mutational analysis of thirty-two double-strand DNA break repair genes in breast and pancreatic cancers. *Cancer Res.*, **68** (4), 971–975.

31. Brandalize, A.P.C., Schüler-Faccini, L., Hoffmann, J.-S., Caleffi, M., Cazaux, C., and Ashton-Prolla, P. (2014) A DNA repair variant in POLQ (c.-1060A > G) is associated to hereditary breast cancer patients: a case-control study. *BMC Cancer*, **14**, 850.

32. Family, L., Bensen, J.T., Troester, M.A., Wu, M.C., Anders, C.K., and Olshan, A.F. (2015) Single-nucleotide polymorphisms in DNA bypass polymerase genes and association with breast cancer and breast cancer subtypes among African Americans and Whites. *Breast Cancer Res. Treat.*, **149** (1), 181–190.

33. Lu, H.-M., Li, S., Black, M.H., Lee, S., Hoiness, R., Wu, S., Mu, W., Huether, R., Chen, J., Sridhar, S., Tian, Y., McFarland, R., Dolinsky, J., Tippin Davis, B., Mexal, S., Dunlop, C., and Elliott, A. (2018) Association of Breast and Ovarian Cancers With Predisposition Genes Identified by Large-Scale Sequencing. *JAMA Oncology*.

34. Kurian, A.W., Hughes, E., Handorf, E.A., Gutin, A., Allen, B., Hartman, A.-R., and Hall, M.J. (2017) Breast and Ovarian Cancer Penetrance Estimates Derived From Germline Multiple-Gene Sequencing Results in Women. *JCO Precision Oncology*, (1), 1–12.

35. Connor, A.A. (2017) Germline & somatic variation in gastrointestinal malignancies.

36. Penkert, J., Schmidt, G., Hofmann, W., Schubert, S., Schieck, M., Auber, B., Ripperger, T., Hackmann, K., Sturm, M., Prokisch, H., Hille-Betz, U., Mark, D., Illig, T., Schlegelberger, B., and Steinemann, D. (2018) Breast cancer patients suggestive of Li-Fraumeni syndrome: mutational spectrum, candidate genes, and unexplained heredity. *Breast Cancer Research*, **20** (1).

37. El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., Qureshi, M., Richardson, L.J., Salazar, G.A., Smart, A., Sonnhammer, E.L.L., Hirsh, L., Paladin, L., Piovesan, D., Tosatto, S.C.E., and Finn, R.D. (2019) The Pfam protein families database in 2019. *Nucleic Acids Research*, **47** (D1), D427–D432.

38. Song, H., Dicks, E., Ramus, S.J., Tyrer, J.P., Intermaggio, M.P., Hayward, J., Edlund, C.K., Conti, D., Harrington, P., Fraser, L., Philpott, S., Anderson, C., Rosenthal, A., Gentry-Maharaj, A., Bowtell, D.D., Alsop, K., Cicek, M.S., Cunningham, J.M., Fridley, B.L., Alsop, J., Jimenez-Linan, M., Høgdall, E., Høgdall, C.K., Jensen, A., Kjaer, S.K., Lubiński, J., Huzarski, T., Jakubowska, A., Gronwald, J., Poblete, S., Lele, S., Sucheston-Campbell, L., Moysich, K.B., Odunsi, K., Goode, E.L., Menon, U., Jacobs,

I.J., Gayther, S.A., and Pharoah, P.D.P. (2015) Contribution of Germline Mutations in the RAD51B, RAD51C, and RAD51D Genes to Ovarian Cancer in the Population. *J. Clin. Oncol.*, **33** (26), 2901–2907.

39. Fan, C., Zhang, J., Ouyang, T., Li, J., Wang, T., Fan, Z., Fan, T., Lin, B., and Xie, Y. (2018) RAD50 germline mutations are associated with poor survival in BRCA1/2-negative breast cancer patients. *Int. J. Cancer*, **143** (8), 1935–1942.

FIGURE LEGENDS

Figure 1 – Variants selection workflow. Whole-exome sequencing data from 52 unrelated Brazilian women at-risk for HBOC, without germline pathogenic variants in *BRCA1*, *BRCA2* and *TP53* genes. Variants classified as "high-impact" and "medium-impact" by snpEFF/GEMINI were prioritized. Then, variants with base coverage ≥10x and variant allele frequency (VAF) ≥0.25 were selected, and those present in population databases with frequency ≤2% (MAF ≤0.02) were analyzed. The variants were also separated accordingly to ClinVar classification.

Figure 2 – Total number of variants per patient. In gray: total number of variants in cancer-associated genes, before applying the ClinVar classification (pathogenic/likely-pathogenic, variants with conflicting interpretations and benign/likely-benign) and *in silico* tools (CADD, REVEL and M-CAP); in black: total number of variants classified as LoF with conflicting interpretations, VUS or not classified by ClinVar; in green: total number of rare missense variants with conflicting interpretations, VUS or not classified by ClinVar and predicted as pathogenic by *in silico* analysis (CADD ≥20, REVEL ≥0.5 and M-CAP ≥0.025).

Figure 3 – Rare variants identified involving genes classified as hallmarks of cancer by COSMIC database. In blue: frameshift variants with conflicting interpretations, VUS or not classified by ClinVar; in yellow: nonsense variants with conflicting interpretations, VUS or not classified by ClinVar; in purple: splice acceptor/donor variants with conflicting interpretations, VUS or not classified by ClinVar; in green: missense variants

with conflicting interpretations, VUS or not classified by ClinVar and predicted as pathogenic by *in silico* analysis (CADD ≥20, REVEL ≥0.5 and M-CAP ≥0.025). Information about tumor diagnosis (in light pink: breast cancer; in dark pink: bilateral breast cancer; in greenish-blue: ovarian cancer; in light blue: melanoma; in black: thyroid cancer) and age at diagnosis (in fluorescent green: diagnosis ≤ 30 years of age; in light yellow: 31 to 45 years of age; in purple: ≥ 46 years of age) are represented.

SUPPLEMENTARY FIGURES

Figure S1 - Variants identified in genes associated with cancer together with information regarding the tumor diagnosed and the patient´s age at diagnosis.

Figure S2 – Total number of rare missense variants identified after applying three *in silico* tools (CADD, REVEL and M-CAP).

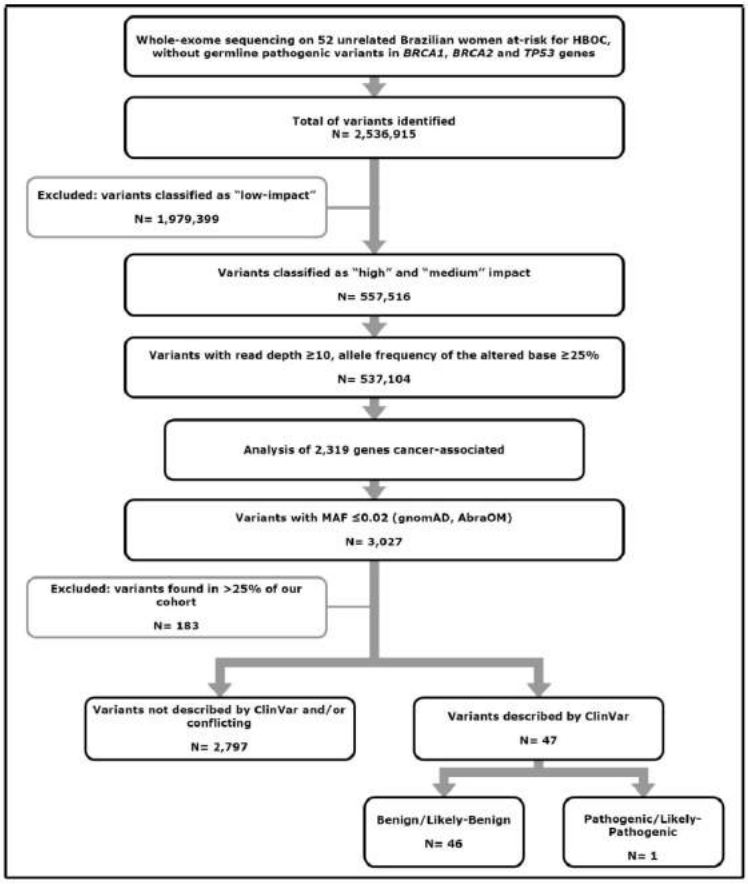


Figure 1 - Variants selection workflow. Whole-exome sequencing data from 52 unrelated Brazilian women at-risk for HBOC, without germline pathogenic variants in BRCA1, BRCA2 and TP53 genes. Variants classified as "high-impact" and "medium-impact" by snpEFF/GEMINI were prioritized. Then, variants with base coverage ≥10x and variant allele frequency (VAF) ≥0.25 were selected, and those present in population databases with frequency ≤2% (MAF ≤0.02) were analyzed. The variants were also separated accordingly to ClinVar classification.

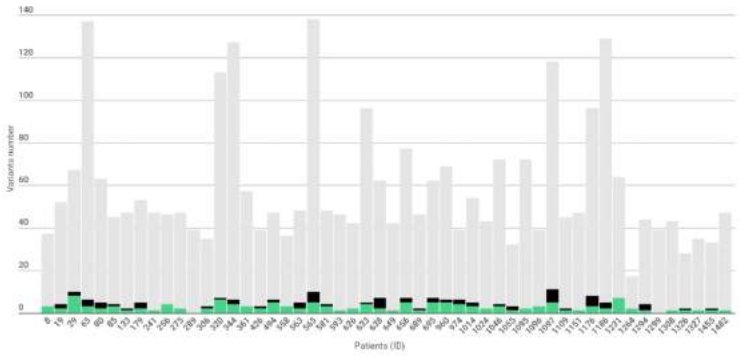


Figure 2 - Total number of variants per patient. In gray: total number of variants in cancer-associated genes, before applying the ClinVar classification (pathogenic/likely-pathogenic, variants with conflicting interpretations and benign/likely-benign) and in silico tools (CADD, REVEL and M-CAP); in black: total number of variants classified as LoF with conflicting interpretations, VUS or not classified by ClinVar; in green: total number of rare missense variants with conflicting interpretations, VUS or not classified by ClinVar and predicted as pathogenic by in silico analysis (CADD ≥20, REVEL ≥0.5 and M-CAP ≥0.025).

658x312mm (72 x 72 DPI)

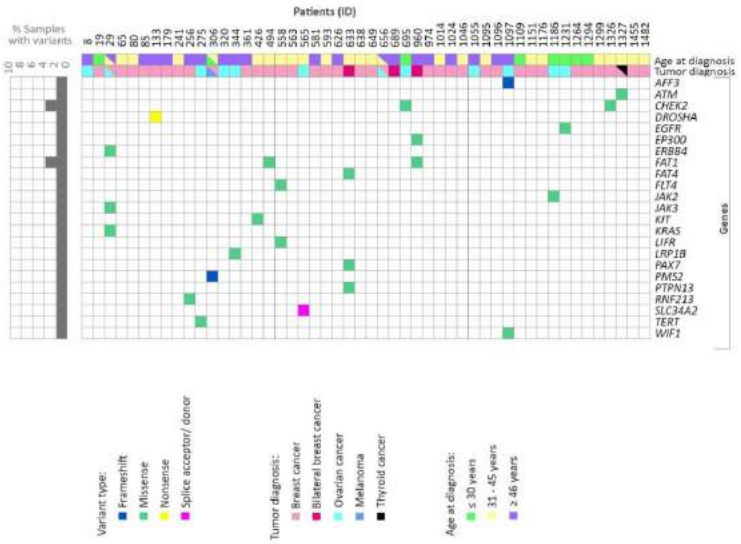


Figure 3 – Rare variants identified involving genes classified as hallmarks of cancer by COSMIC database. In blue: frameshift variants with conflicting interpretations, VUS or not classified by ClinVar; in yellow: nonsense variants with conflicting interpretations, VUS or not classified by ClinVar; in purple: splice acceptor/donor variants with conflicting interpretations, VUS or not classified by ClinVar; in green: missense variants with conflicting interpretations, VUS or not classified by ClinVar and predicted as pathogenic by in silico analysis (CADD ≥20, REVEL ≥0.5 and M-CAP ≥0.025). Information about tumor diagnosis (in light pink: breast cancer; in dark pink: bilateral breast cancer; in greenish-blue: ovarian cancer; in light blue: melanoma; in black: thyroid cancer) and age at diagnosis (in fluorescent green: diagnosis ≤ 30 years of age; in light yellow: 31 to 45 years of age; in purple: ≥ 46 years of age) are represented.

ANEXO IV – Artigo publicado no periódico "*Frontiers in Genetics*" – Produção adicional.

frontiers in Genetics

ORIGINAL RESEARCH
published: 07 May 2018
doi: 10.3389/fgene.2018.00161

Complex Landscape of Germline Variants in Brazilian Patients With Hereditary and Early Onset Breast Cancer

Giovana T. Torrezan[1,2†], Fernanda G. dos Santos R. de Almeida[1†], Márcia C. P. Figueiredo[1], Bruna D. de Figueiredo Barros[1], Cláudia A. A. de Paula[1], Renan Valieris[3], Jorge E. S. de Souza[4,5,6], Rodrigo F. Ramalho[1], Felipe C. C. da Silva[1], Elisa N. Ferreira[1,7], Amanda F. de Nóbrega[6], Paula S. Felicio[9], Maria I. Achatz[8,10], Sandro J. de Souza[2,6,11], Edenir I. Palmero[9,12] and Dirce M. Carraro[1,2]*

[1] Laboratory of Genomics and Molecular Biology, International Research Center, CIPE/A.C. Camargo Cancer Center, São Paulo, Brazil, [2] National Institute for Science and Technology in Oncogenomics and Therapeutic Innovation, São Paulo, Brazil, [3] Laboratory of Bioinformatics and Computational Biology, International Research Center, CIPE/A.C. Camargo Cancer Center, São Paulo, Brazil, [4] Instituto de Bioinformática e Biotecnologia–i2bio, Natal, Brazil, [5] Instituto Metrópole Digital, Federal University of Rio Grande do Norte, Natal, Brazil, [6] Bioinformatics Multidisciplinary Environment, Federal University of Rio Grande do Norte, Natal, Brazil, [7] Research and Development, Fleury Group, São Paulo, Brazil, [8] Oncogenetics Department, A.C. Camargo Cancer Center, São Paulo, Brazil, [9] Molecular Oncology Research Center, Barretos Cancer Hospital, São Paulo, Brazil, [10] Clinical Genetics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, MD, United States, [11] Brain Institute, Federal University of Rio Grande do Norte, Natal, Brazil, [12] Barretos School of Health Sciences, Dr. Paulo Prata – FACISB, Barretos, Brazil*

OPEN ACCESS

Edited by:
Luis G. Carvajal-Carmona, University of California, Davis, United States

Reviewed by:
Tracy A. O'Mara, QIMR Berghofer Medical Research Institute, Australia
John Frederick Pearson, University of Otago, New Zealand

***Correspondence:**
Dirce M. Carraro
dirce.carraro@accamargo.org.br

†These authors have contributed equally to this work.

Specialty section:
This article was submitted to Cancer Genetics, a section of the journal Frontiers in Genetics

Received: *16 November 2017*
Accepted: *17 April 2018*
Published: *07 May 2018*

Citation:
Torrezan GT, de Almeida FGdSR, Figueiredo MCP, Barros BDdF, de Paula CAA, Valieris R, de Souza JES, Ramalho RF, da Silva FCC, Ferreira EN, de Nóbrega AF, Felicio PS, Achatz MI, de Souza SJ, Palmero EI and Carraro DM (2018) Complex Landscape of Germline Variants in Brazilian Patients With Hereditary and Early Onset Breast Cancer. Front. Genet. 9:161. doi: 10.3389/fgene.2018.00161

Pathogenic variants in known breast cancer (BC) predisposing genes explain only about 30% of Hereditary Breast Cancer (HBC) cases, whereas the underlying genetic factors for most families remain unknown. Here, we used whole-exome sequencing (WES) to identify genetic variants associated to HBC in 17 patients of Brazil with familial BC and negative for causal variants in major BC risk genes (*BRCA1/2*, *TP53*, and *CHEK2* c.1100delC). First, we searched for rare variants in 27 known HBC genes and identified two patients harboring truncating pathogenic variants in *ATM* and *BARD1*. For the remaining 15 negative patients, we found a substantial vast number of rare genetic variants. Thus, for selecting the most promising variants we used functional-based variant prioritization, followed by NGS validation, analysis in a control group, cosegregation analysis in one family and comparison with previous WES studies, shrinking our list to 23 novel BC candidate genes, which were evaluated in an independent cohort of 42 high-risk BC patients. Rare and possibly damaging variants were identified in 12 candidate genes in this cohort, including variants in DNA repair genes (*ERCC1* and *SXL4*) and other cancer-related genes (*NOTCH2*, *ERBB2*, *MST1R*, and *RAF1*). Overall, this is the first WES study applied for identifying novel genes associated to HBC in Brazilian patients, in which we provide a set of putative BC predisposing genes. We also underpin the value of using WES for assessing the complex landscape of HBC susceptibility, especially in less characterized populations.

Keywords: cancer predisposition genes, hereditary breast cancer, whole-exome sequencing, germline pathogenic variants, cancer susceptibility, DNA repair genes

INTRODUCTION

Hereditary breast cancer (HBC) corresponds to ~5–10% of all breast cancer cases (Honrado et al., 2005). The most common breast cancer predisposing syndrome is hereditary breast and ovarian cancer syndrome (HBOC) that is related to pathogenic germline variants in *BRCA1* (OMIM 113705) and *BRCA2* (OMIM 600185) genes (Anglian Breast Cancer Study, 2000). These genes correspond to ~20–25% of all HBC (Anglian Breast Cancer Study, 2000; Kean, 2014; Silva et al., 2014). Besides *BRCA1/2* genes, pathogenic variants in other high- and moderate-risk genes, such as *TP53*, *CHEK2*, *ATM*, *STK11*, *PALB2*, among others, also lead to an increased breast cancer (BC) risk, revealing a high complexity in breast cancer predisposition (Elledge and Allred, 1998; Meijers-Heijboer et al., 2002; Walsh and King, 2007).

To date, over 35 genes have been suggested to carry high and/or moderate BC risk variants (OMIM, 2015[1]; Shiovitz and Korde, 2015). However, only a minority of these genes have an established significant association demonstrated by both stringent burden testing and statistical analyses (Easton et al., 2015). Moreover, despite extensive sequencing efforts, variants in known BC susceptibility genes are present in < 30% of BC cases with positive family history or an early age of onset (Shiovitz and Korde, 2015; Chandler et al., 2016), meaning that the underlying genetic factors for most HBC remain unknown.

In the past few years, advances in next-generation sequencing (NGS), specially whole-exome sequencing (WES), have led to the identification of causative variants in several rare familial syndromes, including hereditary cancer (Comino-Méndez et al., 2011; Seguí et al., 2015). Up to the present time, more than 16 different WES studies (both family-based and case studies) have been carried out for HBC, and a few novel BC susceptibility genes were identified: *XRCC2, RINT1, RECQL*, and *FANCM* (Chandler et al., 2016). Nevertheless, the small number of novel major BC autosomal dominant predisposing genes disclosed in these studies has pointed to the possible existence of very rare, or even particular, high and moderate penetrant variants. Conversely, other forms of inheritance, such as recessive and oligogenic transmission of cancer predisposition, cannot be discarded (Sokolenko et al., 2015). In this sense, further WES investigation in different families or populations is crucial for expanding the catalog of breast tumor predisposing genes.

In two previous studies of our group, we screened young BC women (Carraro et al., 2013) and women with clinical criteria of HBOC (Silva et al., 2014) for pathogenic variants in the complete coding sequence of *BRCA1*, *BRCA2*, and *TP53* genes, and for *CHEK2* c.1100delC point mutation, detecting 22–26% of pathogenic variant carriers. Both studies disclosed a large number of women negative for pathogenic variants in the most important genes associated with BC risk, claiming for the necessity of identifying rare and/or novel BC predisposing genes. Thus, the aim of the current study was to investigate, by WES, breast cancer patients with clinical criteria for HBOC and without pathogenic variants in major breast cancer predisposing genes, using rigorous functional criteria for selection of detected variants, in order to identify the most promising new HBC-causing genes.

[1]Online Mendelian Inheritance in Man, OMIM®. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD), 2018. Available online at: https://omim.org/

MATERIALS AND METHODS

Patients and Controls

WES was performed in 17 patients from A.C. Camargo Cancer Center (15 unrelated patients and two siblings) diagnosed with BC and fulfilling one or more of the following criteria of HBOC syndrome: early onset BC (<36 years); bilateral BC; breast plus another primary related tumor (ovary, fallopian tube or primary peritoneal tumors). These patients were selected from previous studies (Carraro et al., 2013; Silva et al., 2014) from our group and were negative for pathogenic variants in *BRCA1/2, TP53*, and *CHEK2* c.1100delC. Two patients (including the two sisters) were carriers of variants of uncertain clinical significance (VUS) in *BRCA1* gene. The detailed inclusion criteria from both studies were described previously (Carraro et al., 2013; Silva et al., 2014). One affected woman of one family participated in the cosegregation study for specific candidate variants.

Five germline *BRCA1*-mutation carriers that were submitted to WES in the same platform were included for variant filtering. For validation of selected variants, target NGS validation was applied in 25 healthy women without family history of cancer, considered here as a control group. Additionally, a selected number of candidate genes were screened in an independent group of 42 patients at risk for HBC from a distinct project, obtained from Barretos Cancer Hospital (Barretos, São Paulo, Brazil). **Figure 1** depicts the study design and workflow, describing the projects steps and the analysis performed in each patients and controls groups.

All participants signed an informed consent. This study was performed in accordance with the Helsinki Declaration and was approved by the A.C. Camargo Cancer Center (1754/13) and Barretos Hospital (916/2015) ethics committees.

DNA Isolation

Genomic DNA was obtained from A.C. Camargo Cancer Center Biobank. In brief, DNA was extracted from peripheral leukocytes by Puregene®-DNA purification Kit (Qiagen, Hilden, Germany), according to manufacturer's instructions. DNA concentration, purity and integrity were assessed by spectrophotometry (Nanodrop 2000—Thermo Fisher Scientific, Waltham, MA) and fluorometry (Qubit—Life Technologies, Foster City, CA, USA).

Whole Exome Sequencing

For the 17 patients of the discovery set, WES was performed using the SOLiD and/or Ion Proton platforms. For SOLiD exomes, libraries were prepared using SOLiD™ Fragment Library Barcoding Kit (Life Technologies) and SureSelect Human All Exon V4 Kit 50 Mb (Agilent Technologies), according to the manufacturer's instructions. Sequencing of paired-end libraries (50 X 75 bp) was performed in a Solid 5500XL System (Life Technologies). For Ion Proton exomes, libraries were prepared using Ion Xpress™ Plus Fragment Library Kit and Ion TargetSeq™ Exome Kit (Thermo Fisher Scientific),

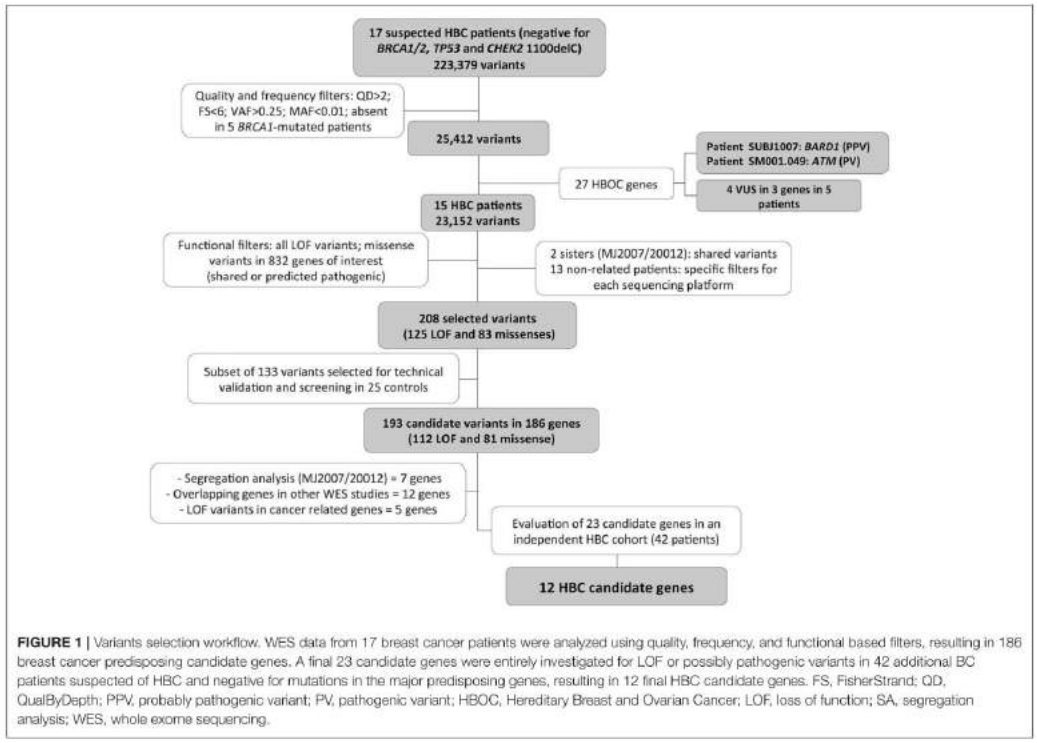


FIGURE 1 | Variants selection workflow. WES data from 17 breast cancer patients were analyzed using quality, frequency, and functional based filters, resulting in 186 breast cancer predisposing candidate genes. A final 23 candidate genes were entirely investigated for LOF or possibly pathogenic variants in 42 additional BC patients suspected of HBC and negative for mutations in the major predisposing genes, resulting in 12 final HBC candidate genes. FS, FisherStrand; QD, QualByDepth; PPV, probably pathogenic variant; PV, pathogenic variant; HBOC, Hereditary Breast and Ovarian Cancer; LOF, loss of function; SA, segregation analysis; WES, whole exome sequencing.

according to the manufacturer's instructions. Each Ion Proton exome library was sequenced on Ion Proton instrument using Ion PI Sequencing 200 Kit v3 and Ion PI Chip v3 (Thermo Fisher Scientific). The resulting sequences were mapped to the reference genome (GRCh37/hg19). Base Calling and alignment were performed by SOLiD™ BioScope 1.2™ Software (Life Technologies) (SOLID data) and by Torrent Suite v4.2 server (Ion Proton data). Variant calling and annotation were done by GATK (Genome Analysis Toolkit) pipeline made available by the Broad Institute. The data obtained in this study is available at Sequence Read Archive (SRP120031).

Variants Selection and Prioritization

For variant filtering, identified variants were annotated with VarSeq (Golden Helix) against reference databases (RefSeq, 1000Genomes, ESP6500, ExAC, dbSNP, and ClinVar). First, for quality filtering, we selected variants with QD > 2 (QD = variant call confidence normalized by depth of sample reads supporting a variant), FS < 6 (FS = strand bias estimated by GATK using Fisher's Exact Test), base coverage ≥ 10x, variant allele frequency (VAF) > 0.25. For four patients with data from both Solid and Ion Proton, only variants detected in both platforms were selected. For one patient with data exclusively from Ion Proton, variants occurring in regions of homopolymer > 4 bases were excluded. Qualified variants were excluded if present in five *BRCA1*-mutation carriers patients analyzed by WES in Solid 5500, and variants present in population databases with frequency > 1% (minor allele frequency [MAF] > 0.01), as well as variants present in more than three unrelated patients. Finally, a recently public available Brazilian database of WES from 609 healthy individuals (Abraom—Brazilian genomic variants; http://abraom.ib.usp.br/) was also used for manually excluding population-specific variants (MAF > 0.01).

Next, for a function-based prioritization, we selected variants leading to loss of function in any gene (frameshift indels, stop codon, and canonical splice site variants) and missense or in-frame indels variants in 832 genes of interest. These genes were selected from commercial panels targeting somatic and germline cancer mutated genes, consensus cancer genes previously described (Futreal et al., 2004) and genes from DNA repair pathways (from KEGG and Putnam et al., 2016) (Supplementary Table 1). For the two related patients, any shared missense or in-frame indels variants in these 832 genes were selected. For the 15 unrelated patients, we selected only variants predicted to be damaging in at least four out of six variant effect prediction software. For these analyses, the results from the following tools were obtained using VarSeq: SIFT, Polyphen v2, Functional Analysis through Hidden Markov

Models (FATHAMM and FATHAMM-MKL), MutationAssessor and MutationTaster. Additionally, we analyzed the potential effect on splicing of the selected LOF and missense variants using dbscSNV annotations (cut-off > 0.6 in ADA and/or RF scores).

Sanger Validation

Two pathogenic variants (PV) or probably pathogenic variants (PPV) in *BARD1* and *ATM* were validated by Sanger sequencing. Briefly, 50 ng of leukocyte DNA was submitted to PCR performed with GoTaq Green Master Mix (Promega), cleaned with ExoSAP-IT (USB Corporation) and sequenced in both directions with BigDye Terminator v3.1 (Life Technologies) using an ABI 3130xl DNA sequencer (Life Technologies), according to manufacturer's instructions. The sequencing results were aligned using CLCBio Genomics Workbench Software (CLCBio, Qiagen). Primer sequences are available under request.

Targeted NGS Validation

A subset of 139 variants (Supplementary Table 2) selected from exome data were validated by multiplex targeted NGS using a custom Ion AmpliSeq panel. Primers were designed using Ion AmpliSeq Designer v3.0.1 (Life Technologies). Libraries were prepared with 20 ng of DNA from each patient using Ion AmpliSeq™ Library Kit 2.0 (Life Technologies). Sequencing was performed using either Ion PGM or Ion Proton platforms, according to the manufacturer's instructions. Sequencing reads mapped to the human genome reference (hg19) using Torrent Suite Browser 4.0.1. On average 166,697 mapped reads were obtained per sample, yielding a mean targeted base coverage of 156X (ranging from 54 to 450). Variants were identified using the VariantCaller v4.0.r73742 plugin and confirmed using CLC Genomics Workbench software (Qiagen). The identified variants were considered if base coverage was ≥10x and VAF > 25%.

To filter out genetic variants common in Brazilian population, the validated variants were evaluated in control group of 25 healthy women by using the same panel. For that, pools of five equimolar genomic DNA samples were prepared by containing 4 ng of each patient (five patients per pool). Libraries preparation, sequencing and mapping were performed as described above. On average 928,194 mapped reads were obtained per pool (mean targeted base coverage 1114X; ranging from 990 to 1,314). Variant calls were obtained using the VariantCaller v4.0.r73742 plugin applying the following filter parameters: VAF > 2%; variant coverage ≥10X.

Cosegregation Analysis

For one family in which a segregation analysis was feasible, DNA from one additional affected individual was obtained. The cosegregation study of specific variants was performed using the same custom gene panel and protocol described previously or with amplicon based library construction and sequencing in Ion Proton platform.

Independent Cohort Validation

For screening the HBC predisposing candidate genes selected in this study an independent cohort comprised of 42 breast cancer patients at risk for HBC from Barretos Cancer Hospital was used. These samples were analyzed through WES in a parallel study using Nextera Rapid Capture Expanded Exome and NextSeq 500 System (Illumina, San Diego, CA). In these data, we assessed the entire coding regions of the 23 genes disclosed in this study for the presence of rare and possibly pathogenic variants, using the same criteria as in our discovery cohort.

RESULTS

In this study we used WES to disclose variants contributing to BC increased risk in patients fulfilling stringent clinical criteria indicating a genetic predisposition to BC and that were negative for pathogenic variants in four major BC genes (*BRCA1/2*, *TP53*, and *CHEK2* 1100delC). The clinical features and family history of cancer for the 17 selected patients are described in Supplementary Table 3.

For the WES, an average of 46,307,427 sequence reads was obtained for each patient and 75.7% (average) of the target bases were covered by 10 or more reads (Supplementary Table 4). More than 200,000 variants were identified in these patients. To prioritize the identified variants, we applied several filters focusing on quality, frequency and function of the identified alterations. The workflow of the variant prioritization is depicted in **Figure 1** and the details of used filters are described in the Materials and Methods section.

Regarding frequency filters, we excluded variants with a minor allele frequency (MAF) >1% in public databases or those present in five germline *BRCA1*-mutation carriers sequenced in our facility, assuming that these variants represent benign or low-penetrance variants. Following these initial data filtering, 25,412 were identified.

Variants in Moderated and High Penetrance Breast Cancer Genes

Initially, we used WES data to search for rare variants in 27 well-established and emerging HBC predisposing genes (the four previously evaluated genes (*BRCA1/2, TP53*, and *CHEK2* c.1100delC) and 23 additional genes): *ATM, BARD1, BLM, BRCA1, BRCA2, BRIP1, CDH1, CHEK2, FANCC, FANCM, MLH1, MSH2, MUTYH, NBN, NF1, PALB2, PMS2, PTEN, RAD51C, RAD51D, STK11, TP53, FAM175A, MRE11, RAD51B, RECQL*, and *RINT1* (Nielsen et al., 2016). In this analysis, we identified two patients harboring frameshift indel variants (one in *ATM* and one in *BARD1*) and five patients (including the two sisters) with variants of uncertain clinical significance (VUS) (**Table 1**). In three patients (MJ2037 and MJ2007/2012) we confirmed the *BRCA1* VUS previously detected by Sanger sequencing. All variants detected in these genes were classified according to the ACMG guidelines (Richards et al., 2015).

The *ATM* p.(Tyr2334Glnfs*4) variant is described as pathogenic in ClinVar database. The *BARD1* p.(Tyr739Leufs*2) is not described in any database and was classified as probably pathogenic, since it is a rare truncating variant leading to partial loss of the second BRCT domain and the phosphobinding region. These two variants were confirmed by Sanger sequencing

TABLE 1 | Pathogenic and VUS detected in 27 known HBC genes.

Patient	Gene	HGVS nomenclature	*N* of 6 Damaging	dbSNP	MAF (ExAC/ Abraom)	Clinical significance (ClinVar)	Clinical significance (ACMG)
SM001.049	*ATM*	c.7000_7003delTACA; p.(Tyr2334Glnfs*4)	–	rs786203421	ND/ND	Pathogenic	Pathogenic
MJ1007	*BARD1*	c.2215dupT; p.(Tyr739Leufs*2)	–	ND	ND/ND	ND	Probably Pathogenic
MJ2003	*RINT1*	c.961T>A; p.(Phe321Ile)	5 of 6	ND	ND/ND	ND	VUS
MJ2001	*RAD51B*	c.728A>G; p.(Lys243Arg)	4 of 6	rs34594234	0.007/0.005	ND	VUS
MJ2037	*BRCA1*	c.5006C>T p.(Ala1669Val)	5 of 6	ND	ND/ND	ND	VUS
MJ2007/2012[#]	*BRCA1*	c.4963T>C; p.(Ser1655Pro)	6 of 6	ND	ND/ND	ND	VUS

[#]*Sisters; N of 6 Damaging: predictions considered as damaging in 6 pathogenicity predicting software; MAF, minor allele frequency; ND, not described; VUS, variant of unknown clinical significance. RefSeq reference number of transcripts are described at Supplementary Table 2.*

in the proband and, for *ATM*, also in one affected relative (Supplementary Figure 1).

Four rare missense variants identified in our patients were classified as probably damaging by at least four prediction software, and three of them are not described in any population database. Three of them are located in recognized functional domains of the affected proteins: *BRCA1* p.Ala1699Val and p.Ser1655Pro are located at the C-terminal BRCT domain, responsible for BRCA1 interaction with others DNA repair proteins and *RINT1* p.Phe321Ile is located at the functional TIP20 domain.

Candidate Selection for Novel Breast Cancer Predisposing Genes

Next, for the 15 patients without any probable pathogenic variant (excluding *ATM* and *BARD1* mutated patients) we applied a functional-based variant prioritization. Candidate variants were selected according to the predicted impact in the protein function and affected gene, including all loss-of-function variants (nonsense, frameshift indels, and splice site) as well as missense and in-frame indels occurring in a list of 832 cancer-related genes (DNA repair and cancer related genes—Supplementary Table 1). For the two sisters (MJ2007 and MJ2012), all variants shared between the two were selected as candidates. For the 13 unrelated patients, we selected missense variants predicted to be damaging by at least 4 out of 6 prediction software.

After filtering, we obtained a total of 208 variants, including 125 LOF and 83 missenses (Supplementary Table 2). In order to technically validate our variant selection workflow, a subset of these 208 variants (133 out of 208) was submitted for technical validation by targeted NGS in the same WES samples and, of these, 126 were validated (95%) (Supplementary Table 4). Using this same custom panel, we evaluated 25 control samples of healthy Brazilian women without cancer for filtering common polymorphisms in our population. Eight variants were detected in at least one control sample and where then excluded from our candidates list, resulting in 193 candidate variants (118 validated and 75 not evaluated).

For the family of the two affected sisters, one additional affected aunt diagnosed with ovarian cancer at age 45 was available for segregation analysis (**Figure 2**). We analyzed 17 variants that were shared between the two sisters and 8 variants were also present in the aunt, including the VUS variant in *BRCA1* (**Table 2**).

Then, the remaining 186 genes prioritized in our study were compared to candidate genes reported in eight previous WES studies of HBC (Snape et al., 2012; Thompson et al., 2012; Gracia-Aznarez et al., 2013; Hilbers et al., 2013; Kiiski et al., 2014; Wen et al., 2014; Noh et al., 2015; Kim et al., 2017) and 12 common genes were identified, 9 of them presenting LOFs variants in at least one study (**Table 3**). For two genes the same LOF variants were identified in our and at a second study (*PZP* p.Arg680* and *KRT76* p.Glu276*).

Thus, from the 193 final candidate variants, we selected 23 candidate genes of BC predisposition: 7 novel candidate genes segregating in the 3 members of the MJ2007/2012 family (*SLC22A16, ROS1, IL33, PTPRD, ARHGEF12, ERBB2, POLA1*), five cancer-related genes harboring LOF variants (*GALNT3, RAF1, PICALM, KL, ERCC1*) and 12 genes overlapping with candidate genes identified in other studies (*CAPN9, KRT76, PZP, DNAH7, MST1R, LAMB4, NIN, MSH3, SLX4, DDX1, NOTCH2,* and *ROS1*—*ROS1* was also selected in the segregating genes list). The entire coding region of the 23 genes were evaluated in an independent Brazilian cohort.

Assessing 23 Candidate Genes in an Independent Cohort of Patients at Risk for HBC

To select the most promising candidate genes, we analyzed the 23 candidate genes disclosed in our study in an independent cohort of 42 Brazilian women at risk for HBC. These patients were all negative for pathogenic variants in *BRCA1/2, TP53,* and *ATM* genes. In these data, we assessed the entire coding regions

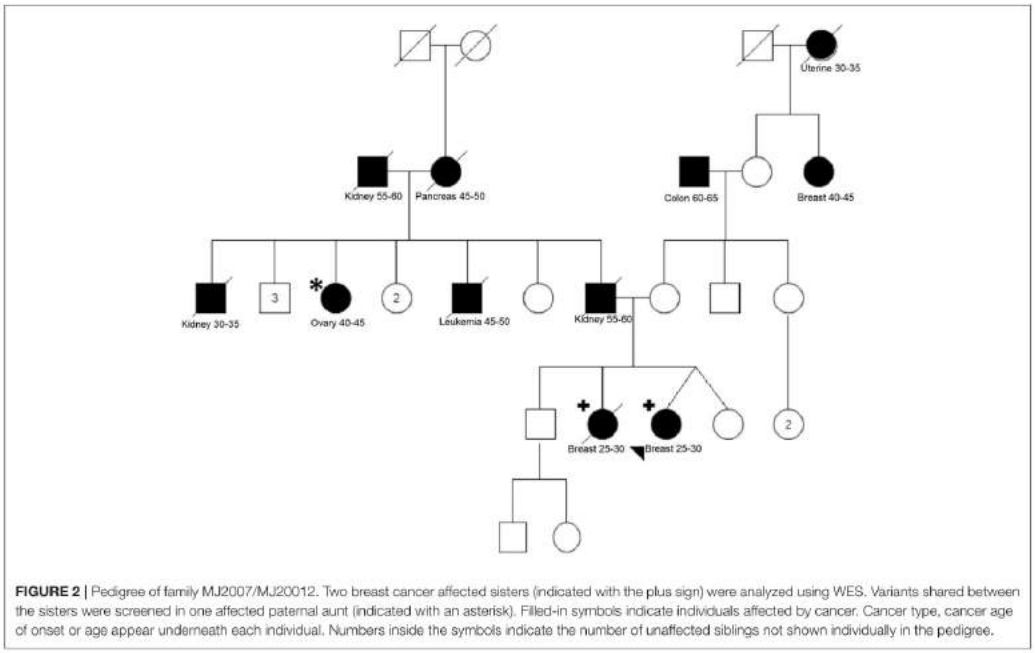


FIGURE 2 | Pedigree of family MJ2007/MJ20012. Two breast cancer affected sisters (indicated with the plus sign) were analyzed using WES. Variants shared between the sisters were screened in one affected paternal aunt (indicated with an asterisk). Filled-in symbols indicate individuals affected by cancer. Cancer type, cancer age of onset or age appear underneath each individual. Numbers inside the symbols indicate the number of unaffected siblings not shown individually in the pedigree.

of the selected genes for the presence of rare (MAF < 1%) and possibly pathogenic variants, selecting all LOF variants and missense variants predicted to be pathogenic in at least 3 out of 6 algorithms.

In this cohort, we detected 16 variants in 12 of the 23 candidate genes (**Table 4**). *NOTCH2* gene was the one with more variants, harboring three missense; *ERBB2* and *DNAH7* harbored two missenses each. Only one LOF variant was detected, affecting *ERCC1* gene, which was the same variant detected in our discovery cohort (c.875G>A; p.Trp292*). The remaining genes presented one rare missense variant each.

DISCUSSION

Recently, the use of WES in clinical genetics has been proven to be an effective alternative for establishing the genetic basis of Mendelian diseases, particularly in diseases where multiple genes can be affected (Trujillano et al., 2016). Moreover, in both clinical and research settings, WES has been applied to elucidate the genetic cause of cancer predisposition. In this sense, WES offers the opportunity to concomitantly investigate several known cancer risk genes as well as to identify novel cancer predisposing genes. Thus, in this study we used WES to disclose variants contributing to BC increased risk in patients that were negative for pathogenic variants in three major BC genes—*BRCA1/2* and *TP53* genes—and the most common point mutation in *CHEK2* gene (c.1100delC). For this, we used stringent clinical criteria for selecting patients with strong indicative of harboring a genetic predisposition to BC, such as early onset BC (<36 years); bilateral BC; or the presence of a second primary related tumor.

First, by evaluating known BC predisposing genes, we could establish the causative variants in two probands. One of them harbored an *ATM* truncating pathogenic variant and the other a novel *BARD1* truncating variant, considered as probably pathogenic. The *BARD1* p.(Tyr739Leufs*2) variant is predicted to cause partial loss of the second functional BRCT domain and the phosphobinding region. Several studies suggest that both BRCT repeats are necessary for BARD1 normal function (Birrane et al., 2007; Irminger-Finger et al., 2016) and truncating variants in this region have been previously reported in association with HBC (De Brakeleer et al., 2010). Additionally, compatible with the probable pathogenic role of this variant, our proband presented triple negative BC and *BARD1* pathogenic variants were recently described to be related to this molecular subtype (De Brakeleer et al., 2016).

Besides these LOF variants, we identified four rare missense VUS in three HBC genes (*BRCA1, RINT1,* and *RAD51B*). The identification of VUS in genetic testing represent a challenging concern for genetic counselors due to uncertainty in clinical decision making, which can lead to more intensive management than necessary in most of the times or, more rarely, in inappropriate prevention measures (Plon et al., 2011). The recently introduction of NGS gene panels in genetic testing

TABLE 2 | Cosegregation analysis of variants detected in the sisters MJ2007 and MJ2012.

Shared variants of MJ2017 and MJ2012								Cosegregation
Chr:Pos	**Ref/Alt**	**Gene names**	**Type**	**HGVS nomenclature**	**N of 6 Damaging**	**dbSNP**	**MAF (ExAC*/ Abraom)**	**Present in affected aunt (OV)**
1:3328745	G/A	*PRDM16*	Missense	c.1984G>A; p.(Val662Met)	1 of 6	ND	ND/0.001	No
1:116670844	G/T	*MAB21L3*	Stop gained	c.739G>T; p.(Glu247*)	2 of 6	rs149122915	0.0002/ND	No
3:48716158	G/A	*NCKIPSD*	Missense	c.1804C>T; p.(Arg602Cys)	2 of 6	ND	0.000008/0.001	No
5:149514363	A/G	*PDGFRB*	Missense	c.581T>C; p.(Ile194Thr)	5 of 6	rs2229560	0.001/0.002	No
6:117622137	C/T	*ROS1*	Missense	c.6733G>A; p.(Gly2245Ser)	2 of 6	rs142264513	0.0008/0.002	No
7:116380062	A/G	*MET*	Missense	c.1451A>G; p.(His484Arg)	2 of 6	ND	0.00005/ND	No
8:17815082	T/G	*PCM1*	Missense	c.1838T>G; p.(Ile613Ser)	0 of 6	rs181777656	0.003 (OT 0.01)/ 0.002	No
14:55467701	T/C	*WDHD1*	Missense	c.703A>G; p.(Ile235Val)	1 of 6	rs139440460	0.004/0.004	No
15:40897315	A/G	*KNL1*	Missense	c.43A>G; p.(Ile15Val)	0 of 6	ND	0.00003/ND	No
6:110778048	C/-	*SLC22A16*	Frameshift	c.226delG; p.(Ala76fs*66)	–	ND	ND/ND	Yes
6:117715381	A/G	*ROS1*	Missense	c.1108T>C; p.(Ser370Pro)	2 of 6	rs56274823	0.002/ND	Yes
9:6255967	G/C	*IL33*	Splice acceptor	c.613-1G>C; p.(spl?)	–	rs146597587	0.002/0.001	Yes
9:8501026	G/A	*PTPRD*	Missense	c.1856C>T; p.(Thr619Ile)	2 of 6	ND	ND/0.001	Yes
11:120298916	C/T	*ARHGEF12*	Missense	c.545C>T; p.(Ser182Phe)	3 of 6	rs147982337	0.002/0.002	Yes
17:37865694	G/A	*ERBB2*	Missense	c.563G>A; p.(Arg188His)	3 of 6	ND	0.00002/0.002	Yes
17:41222968	A/G	*BRCA1*	Missense	c.4963T>C; p.(Ser1655Pro)	6 of 6	ND	ND/ND	Yes
X:24861673	T/C	*POLA1*	Missense	c.3908T>C; p.(Met1303Thr)	2 of 6	ND	ND/ND	Yes

*Chr, chromosome; Pos, position; Ref, reference allele; Alt, alternate allele; N of 6 Damaging, predictions considered as damaging in 6 pathogenicity predicting software; ND, not described; MAF, minor allele frequency; OV, ovary cancer; OT, others. *Variants in ExAc that had a MAF >1% in any ethnic group are underlined and the highest ExAc population MAF is shown inside parenthesis. RefSeq reference number of transcripts are described at Supplementary Table 2.*

have increased the number of patients diagnosed with VUS, emphasizing the urgent need for better pathogenicity predictions models and collaborative efforts to increase observational data that can aid a posteriori classification to variants, such as cosegregation analysis, personal and family history, co-occurrence with pathogenic variants, and histological and molecular features of tumors (Spurdle et al., 2012).

In the 15 patients without known pathogenic variants, we could identify more than 25,000 novel or rare variants (MAF < 1%), thus several filtering strategies were applied to prioritize those more likely to be related to HBC. Since the majority of hereditary cancer predisposing genes harbor an excess of loss-of function variants, we focused on this type of overtly deleterious variants, regardless of the affected gene. Furthermore, most BC risk genes are involved in DNA repair and genomic integrity pathways (Shiovitz and Korde, 2015; Nielsen et al., 2016), and prioritizing variants in these genes is a rational approach that have been used successfully in previous studies (Mantere et al., 2016). As so, we have also focused on missense variants in a defined set of cancer-related and DNA repair genes. By doing that, we were able to reduce our candidate genes list to a few hundreds.

Importantly, for one family with two sisters affected by BC at young ages (29 years), we could improve the selection by retaining only shared variants and also perform segregation analysis of the candidate variants in an aunt affected by ovarian cancer. From this analysis, eight cosegregating variants emerged, including a *BRCA1* VUS. Besides *BRCA1* gene, only *ERBB2* has been previously implicated in BC predisposition, although with conflicting data about the increased risk conferred by some alleles (Breyer et al., 2009; Wang et al., 2013). Regarding the two LOF variants found to be cosegregating in this family (genes *SLC22A16* and *IL33*), no relation between both genes and BC could be recognized in the literature.

One possible explanation for the results observed in this family and that could also be responsible for the cancer predisposition in other patients of our study is the polygenic model. In this model, which has been suggested and reviewed by different authors (Oldenburg et al., 2007; Shiovitz and Korde, 2015), moderate and low penetrance alleles would act in synergy and play a predominant role. Additionally, the high number of affected relatives with different tumor types in both maternal and paternal sides of this family can be a confounding factor for understanding the phenotypes and cosegregation results. Unfortunately, most affected family members of this family were deceased, limiting additional investigations and the interpretation of our findings.

To gain further insight on the relevance of our identified candidate genes, we evaluated the most promising ones in an independent cohort comprising 42 Brazilian HBC women. Several rare and possibly damaging variants were identified in this cohort, providing additional evidence of the potential role in BC predisposition of some new genes. Of those, we highlight four genes related to cancer development and progression (*NOTCH2, ERBB2, MST1R,* and *RAF1*) and two DNA repair genes (*ERCC1* and *SLX4*). Interestingly, *ERCC1* and *SLX4* are partners that act in the repair of interstrand cross-links and are also required for homology-directed repair of DNA double-strand breaks.

TABLE 3 | Overlapping genes selected as candidates in other WES studies.

	15 patients (current study)						Other studies			
Patient	**Gene**	**Type**	**HGVS nomenclature**	**dbSNP**	**MAF (ExAC*/Abraom)**	**Number/type**	**HGVS nomenclature**	**dbSNP**	**MAF (ExAC*/Abraom)**	**Study**
SM001.040	*NOTCH2*	missense	c.2292T>A; p.(Asn764Lys)	ND	0.00002/ND	2 missense	c.854G>A; p.(Arg285His) c.6178C>T; p.(Arg2060Cys)	rs782452794 rs746551843	0.000008/ND 0.000008/ND	Wen et al., 2014
MJ2013S	*CAPN9*	splice site	c.1657+2T>G; p.(spl?)	rs143145032	0.001/0.002	1 frameshift del	c.1976_1982delAGAATG; p.(Glu659Glyfs*20)	ND	0.002/ND	Thompson et al., 2012
MJ2014S and MJ2016S	*DNAH7*	nonsense	c.10359T>G; p.(Tyr3453*)	ND	0.000008/ND	1 frameshift ins	c.4787dupA; p.(Tyr1596*)	rs573013205	0.002/ND	Thompson et al., 2012
SM001.06	*MST1R*	splice site	c.1231-1G>C; p.(spl?)	ND	0.00008/ND	1 nonsense	c.3322C>T; p.(Arg1108*)	rs150876558	0.0001/ND	Thompson et al., 2012
SM001.088	*MSH3*	missense	c.2659G>A; p.(Asp887Asn)	ND	0.00002/ND	1 in frame del	c.162_179del; p.(Ala57_Ala62del)	ND	ND	Thompson et al., 2012
MJ2007/ MJ2012	*ROS1*	missense	c.1100T>C; p.(Ser370Pro)	rs56274923	0.002/ND	1 splice site 1 nonsense	c.5079+2T>C; p.(spl?) c.3303G>A; p.(Trp1101*)	ND rs2001455687	ND 0.00009/ND	Thompson et al., 2012
MJ2004S	*LAMB4*	splice site	c.344-1G>A; p.(spl?)	rs7788865	0.001 (AF 0.0166)/ 0.003	1 frameshift del	c.5265delA; p.(Lys1755Asnfs*11)	rs568834649	0.008/0.004	Thompson et al., 2012
SM001.040	*DDX10*	missense	c.1088G>A; p.(Arg363His)	ND	0.0001/ND	1 nonsense	c.973G>T; p.(Glu325*)	ND	ND	Kiiski et al., 2014
MJ2016S	*PZP*	nonsense	c.2038C>T; p.(Arg680*)	rs145240281	0.005/ND	1 nonsense	c.2038C>T; p.(Arg680*)	rs145240281	0.005 (EU 0.01013)/ ND	Thompson et al., 2012
MJ2015S	*KRT76*	nonsense	c.826G>T; p.(Glu276*)	rs149868801	0.007/0.006	1 nonsense	c.826G>T; p.(Glu276*)	rs149868801	0.007/0.006	Thompson et al., 2012
MJ2014S	*NIN*	in frame del	c.1736_1738delAAG; p.(Glu579del)	ND	0.0003/0.005	1 frameshift del	c.4261delG; p.(Glu1421Lysfs*18)	ND	ND	Thompson et al., 2012
MJ2013S	*SLX4*	missense	c.4766G>A; p.(Arg1589His)	rs746314060	0.00004/0.0008	1 missense	c.2484G>C; p.(Glu828Asp)	rs199656607	0.0001/ND	Kiiski et al., 2014

*ND, not described; MAF, minor allele frequency; AF, Africans; EU, European (Non-Finnish). *Variants in ExAc that had a MAF >1% in any ethnic group are underlined and the highest ExAc population MAF is shown inside parenthesis. RefSeq reference number of transcripts are described at Supplementary Table 2.*

TABLE 4 | Variants detected in 12 candidate genes at the discovery cohort and at an independent cohort.

		15 patients (discovery cohort)					42 patients (validation cohort)				
Patient	**Gene**	**Type**	**HGVS nomenclature**	**N of 6 Damaging**	**dbSNP**	**MAF (ExAC*/ Abraom)**	**Number/ type**	**HGVS nomenclature**	**N of 6 Damaging**	**dbSNP**	**MAF (ExAC/ *Abraom*)**
SM001.040	*NOTCH2*	missense	c.2292T>A; p.(Asn764Lys)	6 of 6	ND	0.00002/ND	3 missense	c.2579T>G; p.(Leu860Trp)	3 of 6	ND	ND/ND
								c.3625T>G; p.(Phe1209Val)	5 of 6	rs147223770	0.003/0.002
								c.7223T>A; p.(Leu2408His)	4 of 6	rs35586704	0.002/0.007
MJ2014S and MJ2016S	*DNAH7*	nonsense	c.10359T>G; p.(Tyr3453*)	–	ND	0.0000008/ND	2 missense	c.3265C>T; p.(Pro1089Ser)	5 of 6	ND	ND/ND
								c.11947C>T; p.(Arg3983Trp)	5 of 6	rs114621989	0.009/0.006
MJ2016S	*RAF1*	frameshift	c.1241dupA; p.(Asp415fs)	–	ND	ND/ND	1 missense	c.923C>T; p.(Pro308Leu)	3 of 6	rs57462220	0.002/0.008
SM001.06	*MST1R*	splice site	c.1231-1G>C; p.(spl?)	–	ND	0.00008/ND	1 missense	c.4180T>A; p.(Ser1394Thr)	4 of 6	rs141336964	0.0002/ND
MJ2004S	*LAMB4*	splice site	c.34+1G>A; p.(spl?)	–	rs77888865	0.001/0.003	1 missense	c.1843C>A; p.(Pro615Thr)	5 of 6	rs201909531	0.0003/0.0008
MJ2014S	*NIN*	in frame del	c.1736_1738delAAG; p.(Glu579del)	–	ND	0.0003/0.005	1 missense	c.848C>T; p.(Ser283Leu)	3 of 6	rs763293400	0.0001/ND
MJ2013S	*SLX4*	missense	c.4766G>A; p.(Arg1589His)	5 of 6	rs74631406	0.00004/0.0008	1 missense	c.3368C>A; p.(Ser1123Tyr)	3 of 6	rs144647122	0.0003/0.0008
MJ2037S	*ERCC1*	nonsense	c.875G>A; p.(Trp292*)	–	rs116640350	0.002 (AF 0.02627)/0.0008	1 nonsense	c.875G>A; p.(Trp292*)	–	rs116640350	0.002/0.0008
MJ2007/2012	*SLC22A16*	frameshift	c.226delG; p.(Ala76fs*66)	–	ND	ND/ND	1 missense	c.599C>T; p.(Ala200Val)	5 of 6	rs617290086	0.003/0.005
MJ2007/2012	*PTPRD*	missense	c.1856C>T; p.(Thr619Ile)	2 of 6	ND	ND/0.0008	1 missense	c.2585G>T; p.(Arg862Leu)	5 of 6	rs142397137	0.0001/ND
MJ2007/2012	*ARHGEF12*	missense	c.545C>T; p.(Ser182Phe)	3 of 6	rs147982337	0.002/0.002	1 missense	c.3986G>A; p.(Arg1329Gln)	3 of 6	ND	ND/ND
MJ2007/2012	*ERBB2*	missense	c.563G>A; p.(Arg188His)	3 of 6	ND	0.00002/0.002	2 missense	c.236A>C; p.(Glu79Ala)	6 of 6	rs617377968	0.0009/0.005
								c.1586T>C; p.(Val529Ala)	3 of 6	ND	ND/ND

*N of 6 Damaging, predictions considered as damaging in 6 pathogenicity predicting software; ND, not described; MAF, minor allele frequency; AF, Africans. *Variants in ExAc that had a MAF > 1% in any ethnic group are underlined and the highest ExAc population MAF is shown inside parenthesis. RefSeq reference number of transcripts are described at Supplementary Table 4.*

Additionally, *ERCC1* is also involved in the nucleotide excision repair pathway (McNeil and Melton, 2012). Both genes have been investigated regarding BC susceptibility, with some common *ERCC1* variants being identified as risk alleles in Chinese population (Yang et al., 2013) and rare truncating and possibly damaging variants in *SLX4* being described in some high risk HBOC patients (Bakker et al., 2013; Shah et al., 2013). Remarkably, in the *ERCC1* gene we identified the same nonsense variant in both discovery and validation cohorts (p.Trp292*), while in *SLX4* one of the rare missense identified in our cohorts (p.Ser1123Tyr) was previously described in one HBC patient (Shah et al., 2013).

Some limitations of our study are inherent to WES method since predisposition variants can be located in non-coding or not captured regions of the genome, such as promoter or deep intronic pathogenic variants. Moreover, although the strategic filtering applied here is necessary to reduce the number of proposed candidates, it can result in the omission of the causative variant (for example, by excluding protein-impacting synonymous variants). Additionally, large genomic rearrangements have been implicated in HBC, and even though specific bioinformatics pipelines can be applied in WES data to extract these results, these analyses were not performed in our study. Finally, when it comes to interpreting the potential effect of our candidate variants in splicing, both coding as well as splice site variants can cause splicing alterations that lead to in-frame functional proteins instead of frameshift truncated ones, and functional assays would be necessary to validate bioinformatics predictions.

Considering the evidence presented here, we can neither conclude that these variants identified in the 15 patients negative for known pathogenic variant are the definitive cause of BC predisposition nor determine the magnitude of the risk that these genes could present. Nevertheless, our results provide a set of novel putative BC predisposing genes and reinforce WES as useful tool for assessing the complex landscape of HBC predisposition. Importantly, this represents the first WES data of a HBC cohort from South America and the analysis of an admixed population such as the Brazilian can reveal unique features compared to other Western populations. In this sense, the WES data generated in our study, as well as other previous and future studies, can be reanalyzed in the future and possibly identify genetic overlaps between families, aiding to gene discoveries (Chandler et al., 2016). Finally, the assignment of a novel gene or specific variant as a true BC predisposition factor requires solid phenotypic evidence from cosegregation analysis, *in vitro* and *in vivo* functional assays and genotyping large series of case and controls from distinct populations. The efforts for discovery and validation of novel HBC genes will continue to provide insights into disease mechanisms, eventually leading to the development of more effective therapies and improved management of affected families.

AUTHOR CONTRIBUTIONS

GT, FdS, EF, and DC: conceived and designed the experiments; GT, FdA, MF, BB, CdP, and EF: performed and analyzed the experiments; RV, JdS, RR, and SdS: performed bioinformatics analysis; AdN, MA, PF, and EP: assessed clinical data, selected, and recruited the patients; SdS, EP, and DC: contributed reagents, materials, and analysis tools; GT, FdA, and DC: wrote and edited the paper. All authors have read and approved the final manuscript.

FUNDING

This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo [2008/57887-9, 2013/23277-8 and 2013/24633-2], Conselho Nacional de Desenvolvimento Científico e Tecnológico [408833/2006-8], FINEP-CT-INFRA (02/2010) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior [23038.004629/2014-19].

ACKNOWLEDGMENTS

We acknowledge the patients and relatives for participating in the study and the A.C. Camargo biobank for sample processing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2018.00161/full#supplementary-material

REFERENCES

Anglian Breast Cancer Study (2000). Prevalence and penetrance of BRCA1 and BRCA2 mutations in a population-based series of breast cancer cases. Anglian Breast Cancer Study Group. *Br. J. Cancer* 83, 1301–1308. doi: 10.1054/bjoc.2000.1407

Bakker, J. L., van Mil, S. E., Crossan, G., Sabbaghian, N., De Leeneer, K., Poppe, B., et al. (2013). Analysis of the novel fanconi anemia gene SLX4/FANCP in familial breast cancer cases. *Hum. Mutat.* 34, 70–73. doi: 10.1002/humu.22206

Birrane, G., Varma, A. K., Soni, A., and Ladias, J. A. (2007). Crystal structure of the BARD1 BRCT domains. *Biochemistry* 46, 7706–7712. doi: 10.1021/bi700323t

Breyer, J. P., Sanders, M. E., Airey, D. C., Cai, Q., Yaspan, B. L., Schuyler, P. A., et al. (2009). Heritable variation of ERBB2 and breast cancer risk. *Cancer Epidemiol. Biomarkers Prev.* 18, 1252–1258. doi: 10.1158/1055-9965.EPI-08-1202

Carraro, D. M., Koike Folgueira, M. A., Garcia Lisboa, B. C., Ribeiro Olivieri, E. H., Vitorino Krepischi, A. C., de Carvalho, A. F., et al. (2013). Comprehensive analysis of BRCA1, BRCA2 and TP53 germline mutation and tumor characterization: a portrait of early-onset breast cancer in Brazil. *PLoS ONE* 8:e57581. doi: 10.1371/journal.pone.0057581

Chandler, M. R., Bilgili, E. P., and Merner, N. D. (2016). A Review of whole-exome sequencing efforts toward hereditary breast cancer susceptibility gene discovery. *Hum. Mutat.* 37, 835–846. doi: 10.1002/humu.23017

Comino-Méndez, I., Gracia-Aznárez, F. J., Schiavi, F., Landa, I., Leandro-García, L. J., Letón, R., et al. (2011). Exome sequencing identifies MAX mutations as a cause of hereditary pheochromocytoma. *Nat. Genet.* 43, 663–667. doi: 10.1038/ng.861

De Brakeleer, S., De Grève, J., Desmedt, C., Joris, S., Sotiriou, C., Piccart, M., et al. (2016). Frequent incidence of BARD1-truncating mutations in germline

DNA from triple-negative breast cancer patients. *Clin. Genet.* 89, 336–340. doi: 10.1111/cge.12620

De Brakeleer, S., De Grève, J., Loris, R., Janin, N., Lissens, W., Sermijn, E., et al. (2010). Cancer predisposing missense and protein truncating BARD1 mutations in non-BRCA1 or BRCA2 breast cancer families. *Hum. Mutat.* 31, 1175–1185. doi: 10.1002/humu.21200

Easton, D. F., Pharoah, P. D., Antoniou, A. C., Tischkowitz, M., Tavtigian, S. V., Nathanson, K. L., et al. (2015). Gene-panel sequencing and the prediction of breast-cancer risk. *N. Engl. J. Med.* 372, 2243–2257. doi: 10.1056/NEJMsr1501341

Elledge, R. M., and Allred, D. C. (1998). Prognostic and predictive value of p53 and p21 in breast cancer. *Breast Cancer Res. Treat.* 52, 79–98. doi: 10.1023/A:1006163101948

Futreal, P. A., Coin, L., Marshall, M., Down, T., Hubbard, T., Wooster, R., et al. (2004). A census of human cancer genes. *Nat. Rev. Cancer* 4, 177–183. doi: 10.1038/nrc1299

Gracia-Aznarez, F. J., Fernandez, V., Pita, G., Peterlongo, P., Dominguez, O., de la Hoya, M., et al. (2013). Whole exome sequencing suggests much of non-BRCA1/BRCA2 familial breast cancer is due to moderate and low penetrance susceptibility alleles. *PLoS ONE* 8:e55681. doi: 10.1371/journal.pone.0055681

Hilbers, F. S., Meijers, C. M., Laros, J. F., van Galen, M., Hoogerbrugge, N., Vasen, H. F. A., et al. (2013). Exome sequencing of germline DNA from non-BRCA1/2 familial breast cancer cases selected on the basis of aCGH tumor profiling. *PLoS ONE* 8:e55734. doi: 10.1371/journal.pone.0055734

Honrado, E., Benítez, J., and Palacios, J. (2005). The molecular pathology of hereditary breast cancer: genetic testing and therapeutic implications. *Mod. Pathol.* 18, 1305–1320. doi: 10.1038/modpathol.3800453

Irminger-Finger, I., Ratajska, M., and Pilyugin, M. (2016). New concepts on BARD1: regulator of BRCA pathways and beyond. *Int. J. Biochem. Cell Biol.* 72, 1–17. doi: 10.1016/j.biocel.2015.12.008

Kean, S. (2014). The " other " breast cancer genes. *Science* 343, 1457–1459. doi: 10.1126/science.343.6178.1457

Kiiski, J. I., Pelttari, L. M., Khan, S., Freysteinsdottir, E. S., Reynisdottir, I., Hart, S. N., et al. (2014). Exome sequencing identifies FANCM as a susceptibility gene for triple-negative breast cancer. *Proc. Natl. Acad. Sci. U.S.A.* 111, 15172–15177. doi: 10.1073/pnas.1407909111

Kim, Y. C., Soliman, A. S., Cui, J., Ramadan, M., Hablas, A., Abouelhoda, M., et al. (2017). Unique features of germline variation in five Egyptian familial breast cancer families revealed by exome sequencing. *PLoS ONE* 12:e0167581. doi: 10.1371/journal.pone.0167581

Mantere, T., Winqvist, R., Kauppila, S., Grip, M., Jukkola-Vuorinen, A., Tervasmäki, A., et al. (2016). Targeted next-generation sequencing identifies a recurrent mutation in MCPH1 associating with hereditary breast cancer susceptibility. *PLoS Genet.* 12:e1005816. doi: 10.1371/journal.pgen.1005816

McNeil, E. M., and Melton, D. W. (2012). DNA repair endonuclease ERCC1-XPF as a novel therapeutic target to overcome chemoresistance in cancer therapy. *Nucleic Acids Res.* 40, 9990–10004. doi: 10.1093/nar/gks818

Meijers-Heijboer, H., van den Ouweland, A., Klijn, J., Wasielewski, M., de Snoo, A., Oldenburg, R., et al. (2002). Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nat. Genet.* 31, 55–59. doi: 10.1038/ng879

Nielsen, F. C., van Overeem Hansen, T., and Sørensen, C. S. (2016). Hereditary breast and ovarian cancer: new genes in confined pathways. *Nat. Rev. Cancer* 16, 599–612. doi: 10.1038/nrc.2016.72

Noh, J. M. , Kim, J., Cho, D. Y., Choi, D. H., Park, W., and Huh, S. J. (2015). Exome sequencing in a breast cancer family without BRCA mutation. *Radiat. Oncol. J.* 33, 149–154. doi: 10.3857/roj.2015.33.2.149

Oldenburg, R. A., Meijers-Heijboer, H., Cornelisse, C. J., and Devilee, P. (2007). Genetic susceptibility for breast cancer: how many more genes to be found? *Crit. Rev. Oncol. Hematol.* 63, 125–149. doi: 10.1016/j.critrevonc.2006.12.004

Plon, S. E., Cooper, H. P., Parks, B., Dhar, S. U., Kelly, P. A., Weinberg, A. D., et al. (2011). Genetic testing and cancer risk management recommendations by physicians for at-risk relatives. *Genet. Med.* 13, 148–154. doi: 10.1097/GIM.0b013e318207f564

Putnam, C. D., Srivatsan, A., Nene, R. V., Martinez, S. L., Clotfelter, S. P., Bell, S. N., et al. (2016). A genetic network that suppresses genome rearrangements in Saccharomyces cerevisiae and contains defects in cancers. *Nat. Commun.* 11256, 1–14. doi: 10.1038/ncomms11256

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* 17, 405–423. doi: 10.1038/gim.2015.30

Seguí, N., Mina, L. B., Lázaro, C., Sanz-Pamplona, R., Pons, T., Navarro, M., et al. (2015). Germline mutations in FAN1 cause hereditary colorectal cancer by impairing DNA repair. *Gastroenterology* 149, 563–566. doi: 10.1053/j.gastro.2015.05.056

Shah, S., Kim, Y., Ostrovnaya, I., Murali, R., Schrader, K. A., Lach, F. P., et al. (2013). Assessment of SLX4 mutations in hereditary breast cancers. *PLoS ONE* 8:e66961. doi: 10.1371/journal.pone.0066961

Shiovitz, S., and Korde, L. A. (2015). Genetics of breast cancer: a topic in evolution. *Ann. Oncol.* 26, 1291–1299. doi: 10.1093/annonc/mdv022

Silva, F. C., Lisboa, B. C. G., Figueiredo, M. C. P., Torrezan, G. T., Santos, É. M. M., Krepischi, A. C., et al. (2014). Hereditary breast and ovarian cancer: assessment of point mutations and copy number variations in Brazilian patients. *BMC Med. Genet.* 15:55. doi: 10.1186/1471-2350-15-55

Snape, K., Ruark, E., Tarpey, P., Renwick, A., Turnbull, C., Seal, S., et al. (2012). Predisposition gene identification in common cancers by exome sequencing: insights from familial breast cancer. *Breast Cancer Res. Treat.* 134, 429–433. doi: 10.1007/s10549-012-2057-x

Sokolenko, A. P., Suspitsin, E. N., Kuligina, E. S., Bizin, I. V., Frishman, D., and Imyanitov, E. N. (2015). Identification of novel hereditary cancer genes by whole exome sequencing. *Cancer Lett.* 369, 274–288. doi: 10.1016/j.canlet.2015.09.014

Spurdle, A. B., Healey, S., Devereau, A., Hogervorst, F. B. L., Monteiro, A. N., Nathanson, K. L., et al. (2012). ENIGMA–evidence-based network for the interpretation of germline mutant alleles: an international initiative to evaluate risk and clinical significance associated with sequence variation in BRCA1 and BRCA2 genes. *Hum. Mutat.* 33, 2–7. doi: 10.1002/humu.21628

Thompson, E. R., Doyle, M. A., Ryland, G. L., Rowley, S. M., Choong, D. Y. H., Tothill, R. W., et al. (2012). Exome sequencing identifies rare deleterious mutations in DNA repair genes FANCC and BLM as potential breast cancer susceptibility alleles. *PLoS Genet.* 8:e1002894. doi: 10.1371/journal.pgen.1002894

Trujillano, D., Bertoli-Avella, A. M., Kumar Kandaswamy, K., Weiss, M. E., Köster, J., Marais, A., et al. (2016). Clinical exome sequencing: results from 2819 samples reflecting 1000 families. *Eur. J. Hum. Genet.* 25, 176–182. doi: 10.1038/ejhg.2016.146

Walsh, T., and King, M.-C. (2007). Ten genes for inherited breast cancer. *Cancer Cell* 11, 103–105. doi: 10.1016/j.ccr.2007.01.010

Wang, H., Liu, L., Lang, Z., Guo, S., Gong, H., Guan, H., et al. (2013). Polymorphisms of ERBB2 and breast cancer risk: a meta-analysis of 26 studies involving 35,088 subjects. *J. Surg. Oncol.* 108, 337–341. doi: 10.1002/jso.23386

Wen, H., Kim, Y. C., Snyder, C., Xiao, F., Fleissner, E. A., Becirovic, D., et al. (2014). Family-specific, novel, deleterious germline variants provide a rich resource to identify genetic predispositions for BRCAx familial breast cancer. *BMC Cancer* 14:470. doi: 10.1186/1471-2407-14-470

Yang, Z., Fang, X., Pei, X., and Li, H. (2013). Polymorphisms in the ERCC1 and XPF genes and risk of breast cancer in a Chinese population. *Genet. Test Mol. Biomarkers* 17, 700–706. doi: 10.1089/gtmb.2013.0122

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ANEXO V - Artigo publicado no periódico "*Asian Pac J Cancer Prev*" – Produção adicional.

DOI:10.31557/APJCP.2019.20.6.1655
Differential Profile of BRCA1 and BRCA2 Mutated Families

RESEARCH ARTICLE

Editorial Process: Submission:04/05/2018 Acceptance:05/31/2019

Differential Profile of *BRCA1* vs. *BRCA2* Mutated Families: A Characterization of the Main Differences and Similarities in Patients

Gabriela Carvalho Fernandes[1], Paula Silva Felicio[1], Rodrigo Augusto Depieri Michelli[2], Aline Silva Coelho[1], Cristovam Scapulatempo-Neto[1,3], Edenir Inêz Palmero[1,2,4]*

Abstract

The identification of families at-risk for hereditary breast cancer (BC) is important because affected individuals present a much higher cancer risk than the general population. The aim of this study was to identify the most important factors associated with the presence of a pathogenic *BRCA1/BRCA2* mutation. Family history (FH), histopathological and immunohistochemical characteristics were compared among BC women with pathogenic *BRCA1/BRCA2* variants; VUSs in B*RCA1/BRCA2*; *BRCA1/BRCA2 WT* and sporadic BC. The most significative differences observed concerned the molecular subtype of the tumors, age at cancer diagnosis and FH of cancer. The presence of bilateral breast cancer (BBC), number of BC cases and the presence of ovarian cancer (OC) increased (respectively) 5.797, 5.033 and 4.412 times the risk of being a *BRCA1/BRCA2* mutation carrier. Besides, women with *BRCA1* or *BRCA2* mutations presented different tumor and FH profiles. The main characteristics associated with a *BRCA1* mutation were triple negativity (OR: 17.31), BBC history (OR: 4.96) and occurrence of OC (OR: 4.32). There were no major discerning components associated with *BRCA2* mutations. Thus, we conclude that tumor pathology and FH of cancer might be considered together at the time of genetic testing mainly in countries where access to genetic testing is still restricted.

Keywords: Hereditary breast cancer- family history of cancer- *BRCA1* vs. *BRCA2* mutated patients

Asian Pac J Cancer Prev, **20 (6),** 1655-1660

Introduction

Cancer has been considered a public health problem for some time in developed and developing countries (UICC, 2016; WHO, 2016). Of the breast cancer (BC) cases diagnosed each year, it is estimated that 5% to 10% are inherited. The main syndrome associated with hereditary BC is known as Hereditary Breast and Ovarian Cancer Predisposition Syndrome (HBOC) (Miki et al., 1994; Wooster et al., 1994), whose main associated genes are *BRCA1* (Miki et al., 1994) and *BRCA2* (Wooster et al., 1994).

It is believed that the *BRCA1/BRCA2* genes are responsible for about 15%-25% of all cases of hereditary breast and ovarian cancer (Easton, 1999; Couch et al., 2014; Mehrgou and Akouchekian, 2016). In addition, studies have reported an increased risk of male breast cancer associated with germline mutations in *BRCA1*, although it represents a lesse frequent association than that with *BRCA2* germline mutations (Struewing et al., 1995; Milne et al., 2008). The *BRCA2* gene increases the risk of developing multiple tumors, such as: tumors of the biliary tract, bladder, esophagus, pancreas, prostate, stomach, melanoma, hematopoietic system, oral cavity and pharynx (Breast Cancer Linkage, 1999; Risch et al., 2006).

In addition to family history, histopathological and immunohistochemical characteristics of tumors are intrinsically related to hereditary breast cancer. Women with triple negative breast cancer (TNBC) who are diagnosed at an early age are candidates for *BRCA1* genetic testing, even if they do not have a family history of breast or ovarian cancer (Lakhani et al., 2002; Oldenburg et al., 2006; Young et al., 2009). On the other side, tumors associated with germline mutations in *BRCA2* generally have immunohistochemical characteristics similar to sporadic tumors (Foulkes et al., 2003; Palacios et al., 2005). Several studies indicate that *BRCA2* tumors are more frequently luminal B subtype (Palacios et al., 2005; Bane et al., 2007; Larsen et al., 2013). Moreover, a study by Bane and colleagues reported that *BRCA2*-associated breast tumors are characterized by an increased expression of fibroblast growth factor 1 and fibroblast growth

[1]*Molecular Oncology Research Center,* [2]*Department of Oncogenetics,* [3]*Department of Pathology, Barretos Cancer Hospital,* [4]*Barretos School of Health Sciences, Dr. Paulo Prata – FACISB, Brazil.* **For Correspondence: edenirip@yahoo.com.br*

factor receptor 2 compared to *BRCA1*-associated breast tumors (Bane et al., 2009). In contrast to breast tumors associated with *BRCA1*, lack of caveolin-1 expression in breast tumors has been reported in patients with germline mutations in *BRCA2*, suggesting that the expression of caveolin-1 occurs only in tumors with mutations in *BRCA1* (Pinilla et al., 2006).

Based on these factors, this study aimed to compare hereditary breast tumors with and without *BRCA1* and *BRCA2* pathogenic variants, with those carrying *BRCA1* and *BRCA2* variants of unknown significance (VUS) and with patients with sporadic breast cancer regarding to the family history of cancer, as well as with histopathological and imunnohistochemical tumor characteristics. In addition, *BRCA1* mutated patients were compared with *BRCA2* mutated ones, in order to identify which are the main similarities and differences between them and which are the most important factors for the identification of a patient with a germline and pathogenic mutation in *BRCA1* or *BRCA2*.

Materials and Methods

Patients

Women with a personal history of breast cancer were included in the study. They were selected and posteriorly classified based on *BRCA1/BRCA2* genetic testing results as follows: Women in Group 1 had a personal and family history of breast cancer with a pathogenic germline variant in *BRCA1* and/or *BRCA2*. Group 2 consisted of women with a personal and family history of breast cancer, with the presence of an identified VUS in *BRCA1* and/or *BRCA2*. Group 3 comprised women with a personal and family history of breast cancer, without a pathogenic variant and/or variant of uncertain clinical significance in *BRCA1* and/or *BRCA2*. Group 4 comprised women with a personal history of breast cancer who were not selected according to their cancer family history and who did not undergo genetic testing for analysis of germline variants in *BRCA1* and/or *BRCA2* ("sporadic"). Women from groups 1, 2, and 3 were recruited through the Oncogenetics Department of the Barretos Cancer Hospital.

The Research Ethics Committee of the Barretos Cancer Hospital approved this project. All participants signed an informed consent form.

Molecular Analysis

For the analysis of mutations in *BRCA1* and *BRCA2*, a multiplex PCR amplification of all coding exons of *BRCA1* (NCBI; NM_007294.3) and *BRCA2* (NCBI; NM_000059.3) and their respective flanking intronic regions was performed. This was followed by bi-directional Sanger sequencing (ABI 3500XL, Applied Biosystems) as described elsewhere (Fernandes et al., 2016; Palmero et al., 2016a)). In addition, large rearrangements were investigated using the multiplex ligation-dependent probe amplification (MLPA) technique.

For the classification of the germline *BRCA1* and *BRCA2* variants identified ClinVar (ClinVar) and *BRCA* share databases(UMD) were used.

Immunohistochemistry

Monoclonal antibodies were used against ER, PR and Ki-67, CK5/6, and CK14, and a polyclonal antibody was used against HER2. For the analysis of staining, ER, PR, CK5/6, and CK14 were considered as either negative or positive. For the cell proliferation marker Ki-67, the indexes were grouped into the categories ≤14% and >14%. For the HER2 receptor, in addition to the positive and negative categories, a third category termed inconclusive was added. In these cases, fluorescent in situ hybridization (FISH) analysis using SPEC Her-2/CEN17 Dual Color Probe were performed.

Clinical and family history data

The clinical data of the patients included in the study were obtained from the patient's general medical records. Family history data were obtained from the medical records of the Oncogenetics Department of the Barretos Cancer Hospital.

Statistical Analysis

The program Statistical Package for Social Sciences v.21.0 for Windows (Chicago, IL) was used for the statistical analysis. Categorical variables were described using absolute frequencies and relative percentage frequencies. Correlations were performed using Chi-square and Fisher's exact tests. Bonferroni correction for multivariate analysis was used to estimate the predictive effects of the significantly associated factors for predicting the probability of *BRCA1/BRCA2* mutations. The level of significance adopted in all tests was 5%.

Results

General Characteristics

Patients included in the study were women with a personal history of breast cancer. Group 1 had 51 patients, Group 2 comprised 53 women, Group 3 comprised 100 women, and Group 4 had 83 women.

The average age at diagnosis was 41.88 years (SD = 10.5 years) in Group 1, 34.91 years (SD = 10.0 years) in Group 2, 38.37 years (SD = 10.8 years) in Group 3, and 51.65 years (SD = 9.9 years) in Group 4. Detailed age distribution can be found in Supplementary Table 1. With regard to hormonal risk factors for breast cancer, we observed that 23.6% of the women were menopausal at diagnosis, and 79.4% had a previous pregnancy. The pathological data of the samples are summarized in Supplementary Table 2. Patients in Group 1 had a higher proportion of histological grade III and T4 tumors when compared to the patients from Groups 2–4. As was expected, most of the women in Group 1 had TNBC. In addition, when Group 1 was stratified according to mutated gene, we observed that, among the triple negative tumors, 91.6% had a pathogenic mutation in *BRCA1*, and only 8.4% in *BRCA2*.

When histological analysis was performed dividing Group 1 by mutated gene, we observed a small difference in histological grade, namely, 56.5% of women with a germline mutation in *BRCA1* had histological grade III tumors, compared to 25.0% of women with a germline

Table 1. Molecular Subtype of Breast Tumors by Group

Variable	Group 1 N (%)		Group 2 N (%)	Group 3 N (%)	Group 4 N (%)	p-value
	MUTATED BRCA1	MUTATED BRCA2				
Molecular subtype						**
Luminal A	1 (7.7)	0 (0.0)	1 (2.6)	8 (11.0)	19 (26.0)	
Luminal B (HER2 negative)	4 (30.8)	12 (70.6)	22 (57.9)	25 (34.2)	31 (42.5)	
Luminal B (HER2 positive)	0 (0.0)	4 (23.5)	9 (23.7)	19 (26.0)	12 (16.4)	
HER2 overexpressed	0 (0.0)	0 (0.0)	5 (13.2)	11 (15.1)	7 (9.6)	
Basal-like	8 (61.5)	1 (5.9)	1 (2.6)	10 (13.7)	4 (5.5)	

**, Given the very low sample size in some of the groups, it was not possible to determine statistical significance.

mutation in *BRCA2* (p = 0.03). The other variables (size tumor, lymph node, metastasis) were also analyzed according to the mutated gene, but no statistically significant difference was found.

The majority of the breast tumors were ductal infiltrating carcinomas (236 cases, 91.8%), followed by lobular carcinomas (10 cases, 3.9%), and medullary carcinomas (4 cases, 1.6%). Medullary carcinomas were further investigated and it was found that two of them belonged to Group 1, one of whom had a mutation in *BRCA1*, and the other in *BRCA2*. The other two patients with medullary carcinoma were from Groups 2 and 4 respectively.

The results of the immunohistochemistry are reported in Supplementary Table 3. A higher positivity of CK5/6 and CK14 staining was observed in Group 1 compared to the other groups, and, once more, this increase in positivity was mainly associated with *BRCA1* pathogenic variants.

The predominant molecular subtype in all of the groups, except for those *BRCA1* mutated inside group 1, was luminal B without HER2 expression. As expected, among the group with sporadic breast cancer, the second most common subtype was luminal A. Moreover, a substantial number (30%) of tumors from Group 1 were classified as basal-like (88.8% of them with a germline mutation in *BRCA1*) (Table 1).

Regarding the expression of the evaluated cytokeratins, the cancers associated with a germline pathogenic variant in *BRCA1* were more frequently positive for CK5/6 and CK14 than tumors associated with a germline pathogenic variant in *BRCA2* (p = 0.031 and p = 0.008, respectively).

Family history of cancer

The pedigree of the 204 families at-risk for hereditary breast cancer included in the study were criteriously revised (Supplementary Table 1). Through the analyzed family history, it was observed that the majority of women in Group 1 reported more than three cases of breast cancer in the family (28 participants, 54.9%), while most of women in the other analyzed groups reported fewer than three cases of breast cancer (p < 0.001). As expected, the presence of bilateral breast cancer was reported most by patients in Group 1 with 10 cases (19.6%), whereas only 5 patients in Group 2 and 5 patients in Group 3 reported the presence of bilateral breast cancer (p < 0.001). As it would be expected, no patients in Group 4 were diagnosed with bilateral breast cancer, male cancer or family history with more than 3 breast cancer cases.

In addition, the presence of ovarian cancer in the family history was observed more frequently in Group 1 (13 cases, 25.5%) than in groups 2 and 3, and no patients in Group 4 reported the presence of ovarian cancer in their families (p < 0.001). It should be noted that patients from Group 4 were not tested for germline mutations in

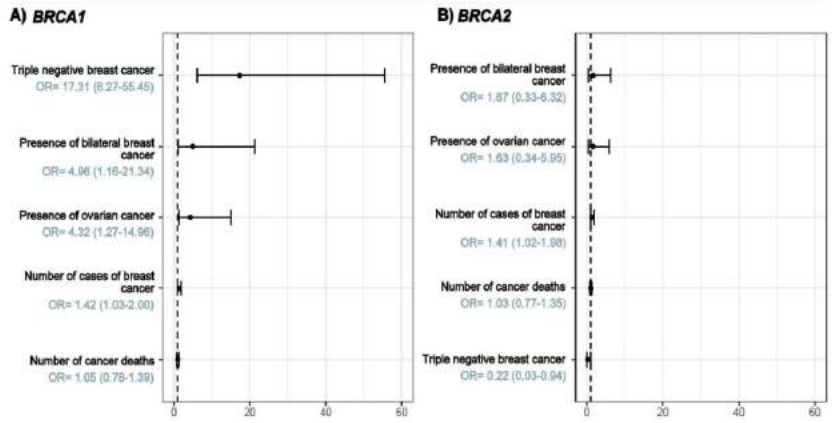


Figure 1. Main Factors Associated with the Presence of a *BRCA1* (in A) or *BRCA2* (in B) Pathogenic Variant

BRCA1/BRCA2 because of the criteria adopted by the Oncogenetics Department of Barretos Cancer Hospital (Palmero et al., 2016b).

Comparison of patients with BRCA1/BRCA2 mutations versus patients with WT BRCA1/BRCA2 and sporadic breast cancer cases regarding histopathological, molecular, and family history

In order to identify which characteristics were typical and representative of the patients with *BRCA1/BRCA2* pathogenic germline variants, a multivariate analysis, in which the *BRCA1/BRCA2* mutation carriers were compared with the other three groups was performed. However, as observed previously, there were major differences between patients with *BRCA1* and *BRCA2* pathogenic variants. For this reason, the comparisons were performed separately for both groups. Detailed results can be found in Supplementary Tables 4 and 5 for *BRCA1* and *BRCA2* carriers, respectively.

In relation to the family history, some variables that showed significance in the compared groups should be highlighted, such as the presence of bilateral breast cancer, ovarian cancer, number of generations affected by cancer, number of breast cancer cases and age at diagnosis. Regarding tumor characteristics, the essential variables that came out are ER, PR, HER2, and cytokeratin expression (Supplementary Tables 3 and 4).

After that, a multivariate analysis – logistic regression was performed and allow the identification of the central characteristics that differed between patients with and without mutations and to see the "weight" that each of these characteristics conferred upon the likelihood of carrying a germline mutation in *BRCA1* or *BRCA2* (Figure 1). Comparing the *BRCA1* (A) and *BRCA2* (B) carriers according to personal/family history, it was possible to see that triple negativity, bilateral breast and ovarian cancer in the proband or family were the main factors associated with the presence of a *BRCA1* mutation. For *BRCA2*, there were no variables in the family history or tumor profile conferring a significative higher risk to be a carrier, highlighting how different the tumors of individuals carrying germline mutations in *BRCA1* and *BRCA2* are, in spite of their association with the same cancer predisposition syndrome, HBOC and which are the main differences and similarities between them.

Discussion

A family history of cancer in first-degree relatives and the presence of some specific risk factors, such as bilateral breast cancer, family history of breast and ovarian cancer, and breast cancer in a male person are important indicators of risk for hereditary breast cancer. Advances in molecular biology in recent decades have resulted in the identification of genes, such as the tumor suppressor genes *BRCA1* and *BRCA2*, that, when altered, significantly increase the risk of developing breast cancer, ovarian cancer, and other tumors. Additionally, especially in developing and undeveloped countries, access to testing is still restricted owing to its high cost and the lack of coverage by health plans; therefore, the correct identification of individuals and families that would benefit from genetic testing is extremely important (Palmero et al., 2016b).

In this study, we included 287 women with breast cancer from Barretos Cancer Hospital. The main goal of this study was to identify the principal characteristics that differentiate a family with *BRCA1/BRCA2* pathogenic mutations from those with a VUS alteration and even from those *BRCA1/BRCA2 WT* and with sporadic breast tumors. For that, clinical data, tumor profile and family history of cancer were compared among the women assigned to each of the four main groups of the study. However, we find out that the variables typical of *BRCA1* carriers could not be applied to identify *BRCA2* carriers.

The most common histological type diagnosed among the women from all four grous was invasive ductal adenocarcinoma (84.4%). Even for women with an identified germline mutation, the presence of medullary carcinoma was low (1.6%). The CIMBA group analyzed 4,325 mutation carriers and found that medullary tumors were 9.4% and 2.2% of tumors identified among *BRCA1* and *BRCA2* mutation carriers (Mavaddat et al., 2012). No difference in histological grade was observed among the four groups of patients. However, it is interesting to point that, in Group 1, a significant number of cases were grade III (43.6%). Moreover, when stratified according to the mutated gene, we observed that the majority of those grade III cases harbored *BRCA1* mutations (56.5% vs. 25.0% *BRCA2* mutations). Although in a minor proportion, our data are in accordance with the results published by the CIMBA group, where the authors reported that 77.0% of women with mutations in *BRCA1* had grade III breast tumors, compared to only 50.0% of women with mutations in *BRCA2* (Mavaddat et al., 2012).

Regarding the hormone receptors (ER, PR), most tumors from *BRCA1* mutated patients in Group 1 were negative for both ER and PR, unlike the other groups and also the *BRCA2* mutated from group 1, where positivity for both receptors prevailed. Most patients, regardless of group, were HER2-negative. When stratifying the women from Group 1 according to the mutated gene, we observed that 88.9% of the women with TNBC were *BRCA1* mutation carriers, while only 11.1% had a *BRCA2* mutation. These data corroborate previous findings in the literature, which show a higher incidence of triple negativity in women with a deleterious *BRCA1* mutation (Bayraktar et al., 2011; Evans et al., 2011; Hartman et al., 2012; Meyer et al., 2012; Triantafyllidou et al., 2015; Wong-Brown et al., 2015). In addition, several studies have identified individuals with *BRCA1/BRCA2* germline mutations through the analysis of TNBC cases. The rate of *BRCA1/BRCA2* mutation detection in those studies, selected based on the tumor triple negativity, independent of family history, varies from 17.4% to 49.1% (Hartman et al., 2012; Couch et al., 2015). Besides, our data on the sporadic group (Group 4) supports the findings from the literature, with a 15% to 20% frequency of TNBC(Bauer et al., 2007; Blows et al., 2010; Lin et al., 2012).

For cases in which the result of immunohistochemistry could be obtained, we conducted a classification of molecular subtype as described by Goldhirsch(Goldhirsch

et al., 2011). The data from this analysis showed that 30% of women with a germline mutation in *BRCA1* or *BRCA2* had a basal-like molecular subtype, which is lower than what has been described in the literature (Andres et al., 2014). However, when analyzing *BRCA1* mutation vs. *BRCA2* mutation, we observed that 88.8% of women with a germline mutation in *BRCA1* displayed basal-like histology versus 11.2% of those with a germline mutation in *BRCA2*. A study conducted by Pinilla et al. also reported an increased frequency of basal-like tumors in patients with *BRCA1* mutations (55.6%) compared to *BRCA2* mutation carriers (10%) (Pinilla et al., 2006).

Similar to the results of other studies, we observed an association between germline mutations in *BRCA1/BRCA2* and a family history of cancer. When comparing the presence/absence of a germline mutation with a family history of cancer, we noted that 54.9% of women in Group 1 had more than three cases of breast cancer in their family history. In addition, the presence of bilateral breast cancer was most reported by patients with a germline mutation (19.6%) compared to women without a germline mutation. This association was observed in other studies, as described by Gershoni-Baruch et al., (1999), in which 45% of patients with a germline mutation (*BRCA1/BRCA2*) and positive family history were diagnosed with bilateral breast cancer (14/31). Regarding the difference found in the family history of women with mutations in *BRCA1* or *BRCA2* genes, we can observe that BRCA1 mutated families showed a higher frequency of bilateral breast cancer (OR=4.96 vs. OR=1.67) and presence of ovarian cancer (OR=4.32 vs. OR=1.63).

It is noteworthy that a family history of cancer is the main indicator for genetic testing. However, several studies have pointed to the fact that many women are candidates for genetic testing by combining tumor characteristics, such as histopathology and immunohistochemistry and factors related to the personal and family history of cancer (Mavaddat et al., 2012; Spurdle et al., 2014). Taking this into consideration, a multivariate analysis was performed to identify the main characteristics associated with the presence of *BRCA1* or *BRCA2* mutations. This analysis allowed us to identify that BRCA1 was mainly correlated with triple negativity (OR: 17.31), presence of bilateral breast cancer history (OR: 4.96), occurrence of ovarian cancer (OR: 4.32) and presence of more than three breast cancer cases in the family (OR: 1.42). When only *BRCA2* mutation carriers were considered, there were no major discerning characteristics. These results reinforce the necessity that family history should be considered, but not in isolation, as a factor of selection and identification of families to be referenced for genetic testing. A more detailed investigation would increase the rate of detection of *BRCA1/BRCA2* mutation carriers, allowing health care providers to direct resources and expand genetic testing access to those families that have higher risk and probability of having a pathogenic *BRCA* mutation, which is very important, particularly in those countries where genetic testing is still restricted.

Finally, it is worth noting that despite the relatively restrictive criteria applied by the Oncogenetics Department for the selection of patients who should be tested and the broad methodology for the analysis of genes involved, the high/moderate risk of cancer attributed to family (in the case of Groups 2 and 3 of this study) remains unexplained. Part of this may be due to the fact that there may be genetic alterations in other genes not yet associated with hereditary breast cancer or the presence of genetic alterations in other high/moderate risk genes for which the patients were not tested (such as *TP53* and *PALB2*). In addition, it emphasizes the importance of characterizing clinical, pathological, and molecular data, including other variables, as well as family history of cancer, as a criterion for the selection/identification of women and families at-risk for hereditary breast cancer. In conclusion, this research combines both pedigree and tumor data to identify the main variables associated with the presence of a *BRCA1* or *BRCA2* germline mutation. Prediction of *BRCA1/BRCA2* carrier status, and hence selection of women for mutation screening, may be substantially improved by combining tumor pathology with family history. These variables must be considered together at the time of genetic counseling mainly in countries where access to genetic testing is still restricted.

Funding

This work was supported by a grant from Conselho Nacional de Desenvolvimento Cientifico e Tecnológico (CNPq), Edital MCT/CNPq 14/2010, process 480760/2010-1. EIP receives a National Council of Technological and Scientific Development (CNPq) scholarship.

Competing interests

The authors declare that they have no competing interests.

Acknowledgments

The authors thanks to the Research Support Nucleo from the Barretos Cancer Hospital.

References

Andres R, Pajares I, Balmana J, et al (2014). Association of *BRCA1* germline mutations in young onset triple-negative breast cancer (TNBC). *Clin Transl Oncol*, **16**, 280-4.

Bane AL, Beck JC, Bleiweiss I, et al (2007). BRCA2 mutation-associated breast cancers exhibit a distinguishing phenotype based on morphology and molecular profiles from tissue microarrays. *Am J Surg Pathol*, **31**, 121-8.

Bane AL, Pinnaduwage D, Colby S, et al (2009). Expression profiling of familial breast cancers demonstrates higher expression of FGFR2 in BRCA2-associated tumors. *Breast Cancer Res Treat*, **117**, 183-91.

Bauer KR, Brown M, Cress RD, et al (2007). Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype: a population-based study from the California cancer Registry. *Cancer*, **109**, 1721-8.

Bayraktar S, Gutierrez-Barrera AM, Liu D, et al (2011). Outcome of triple-negative breast cancer in patients with or without deleterious BRCA mutations. *Breast Cancer Res Treat*, **130**, 145-53.

Blows FM, Driver KE, Schmidt MK, et al (2010). Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies. *PLoS Med*, **7**, e1000279.

Breast Cancer Linkage C (1999). Cancer risks in BRCA2 mutation carriers. *J Natl Cancer Inst*, **91**, 1310-6.

ClinVar. National Center for Biotechnology Information, U.S. National Library of Medicine [Online]. Available: https://www.ncbi.nlm.nih.gov/clinvar/.

Couch FJ, Hart SN, Sharma P, et al (2015). Inherited mutations in 17 breast cancer susceptibility genes among a large triple-negative breast cancer cohort unselected for family history of breast cancer. *J Clin Oncol*, **33**, 304-11.

Couch FJ, Nathanson KL, Offit K (2014). Two decades after BRCA: setting paradigms in personalized cancer care and prevention. *Science*, **343**, 1466-70.

Easton DF (1999). How many more breast cancer predisposition genes are there?. *Breast Cancer Res*, **1**, 14-7.

Evans DG, Howell A, Ward D, et al (2011). Prevalence of *BRCA1* and *BRCA2* mutations in triple negative breast cancer. *J Med Genet*, **48**, 520-2.

Fernandes GC, Michelli RA, Galvao HC, et al (2016). Prevalence of *BRCA1/BRCA2* mutations in a Brazilian population sample at-risk for hereditary breast cancer and characterization of its genetic ancestry. *Oncotarget*, **7**, 80465-81.

Foulkes WD, Stefansson IM, Chappuis PO, et al (2003). Germline *BRCA1* mutations and a basal epithelial phenotype in breast cancer. *J Natl Cancer Inst*, **95**, 1482-5.

Gershoni-Baruch R, Dagan E, Fried G, et al (1999). *BRCA1* and *BRCA2* founder mutations in patients with bilateral breast cancer. *Eur J Hum Genet*, **7**, 833-6.

Goldhirsch A, Wood WC, Coates AS, et al (2011). Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *Ann Oncol*, **22**, 1736-47.

Hartman AR, Kaldate RR, Sailer LM, et al (2012). Prevalence of BRCA mutations in an unselected population of triple-negative breast cancer. *Cancer*, **118**, 2787-95.

Lakhani SR, Van De Vijver MJ, Jacquemier J, et al (2002). The pathology of familial breast cancer: predictive value of immunohistochemical markers estrogen receptor, progesterone receptor, *HER-2*, and *p53* in patients with mutations in *BRCA1* and *BRCA2*. *J Clin Oncol*, **20**, 2310-8.

Larsen MJ, Kruse TA, Tan Q, et al (2013). Classifications within molecular subtypes enables identification of *BRCA1/BRCA2* mutation carriers by RNA tumor profiling. *PLoS One*, **8**, e64268.

Lin NU, Vanderplas A, Hughes ME, et al (2012). Clinicopathologic features, patterns of recurrence, and survival among women with triple-negative breast cancer in the National Comprehensive Cancer Network. *Cancer*, **118**, 5463-72.

Mavaddat N, Barrowdale D, Andrulis IL, et al (2012). Pathology of breast and ovarian cancers among BRCA1 and BRCA2 mutation carriers: results from the Consortium of Investigators of Modifiers of *BRCA1/2 (CIMBA)*. *Cancer Epidemiol Biomarkers Prev*, **21**, 134-47.

Mehrgou A, Akouchekian M (2016). The importance of *BRCA1* and *BRCA2* genes mutations in breast cancer development. *Med J Islam Repub Iran*, **30**, 369.

Meyer P, Landgraf K, Hogel B, et al (2012). BRCA2 mutations and triple-negative breast cancer. *PLoS One*, **7**, e38361.

Miki Y, Swensen J, Shattuck-Eidens D, et al (1994). A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*. *Science*, **266**, 66-71.

Milne RL, Osorio A, Cajal TR, et al (2008). The average cumulative risks of breast and ovarian cancer for carriers of mutations in *BRCA1* and *BRCA2* attending genetic counseling units in Spain. *Clin Cancer Res*, **14**, 2861-9.

Oldenburg RA, Kroeze-Jansema K, Meijers-Heijboer H, et al (2006). Characterization of familial non-*BRCA1/2* breast tumors by loss of heterozygosity and immunophenotyping. *Clin Cancer Res*, **12**, 1693-700.

Palacios J, Honrado E, Osorio A, et al (2005). Phenotypic characterization of *BRCA1* and *BRCA2* tumors based in a tissue microarray study with 37 immunohistochemical markers. *Breast Cancer Res Treat*, **90**, 5-14.

Palmero EI, Alemar B, Schuler-Faccini L, et al (2016a). Screening for germline *BRCA1*, *BRCA2*, *TP53* and *CHEK2* mutations in families at-risk for hereditary breast cancer identified in a population-based study from Southern Brazil. *Genet Mol Biol*, **39**, 210-22.

Palmero EI, Galvao HC, Fernandes GC, et al (2016b). Oncogenetics service and the Brazilian public health system: the experience of a reference Cancer Hospital. *Genet Mol Biol*, **39**, 168-77.

Pinilla SM, Honrado E, Hardisson D, et al (2006). Caveolin-1 expression is associated with a basal-like phenotype in sporadic and hereditary breast cancer. *Breast Cancer Res Treat*, **99**, 85-90.

Risch HA, McLaughlin JR, Cole DE, et al (2006). Population *BRCA1* and *BRCA2* mutation frequencies and cancer penetrances: a kin-cohort study in Ontario, Canada. *J Natl Cancer Inst*, **98**, 1694-706.

Spurdle AB, Couch FJ, Parsons MT, et al (2014). Refined histopathological predictors of *BRCA1* and *BRCA2* mutation status: a large-scale analysis of breast cancer characteristics from the *BCAC*, *CIMBA*, and *ENIGMA* consortia. *Breast Cancer Res*, **16**, 3419.

Struewing JP, Brody LC, Erdos MR, et al (1995). Detection of eight *BRCA1* mutations in 10 breast/ovarian cancer families, including 1 family with male breast cancer. *Am J Hum Genet*, **57**, 1-7.

Triantafyllidou O, Vlachos IS, Apostolou P, et al (2015). Epidemiological and clinicopathological characteristics of BRCA-positive and BRCA-negative breast cancer patients in Greece. *J BUON*, **20**, 978-84.

UICC. 2016. Union for International Cancer Control [Online]. Available: http://www.uicc.org/.

UMD. BRCA Share Database [Online]. Available: http://www.umd.be/.

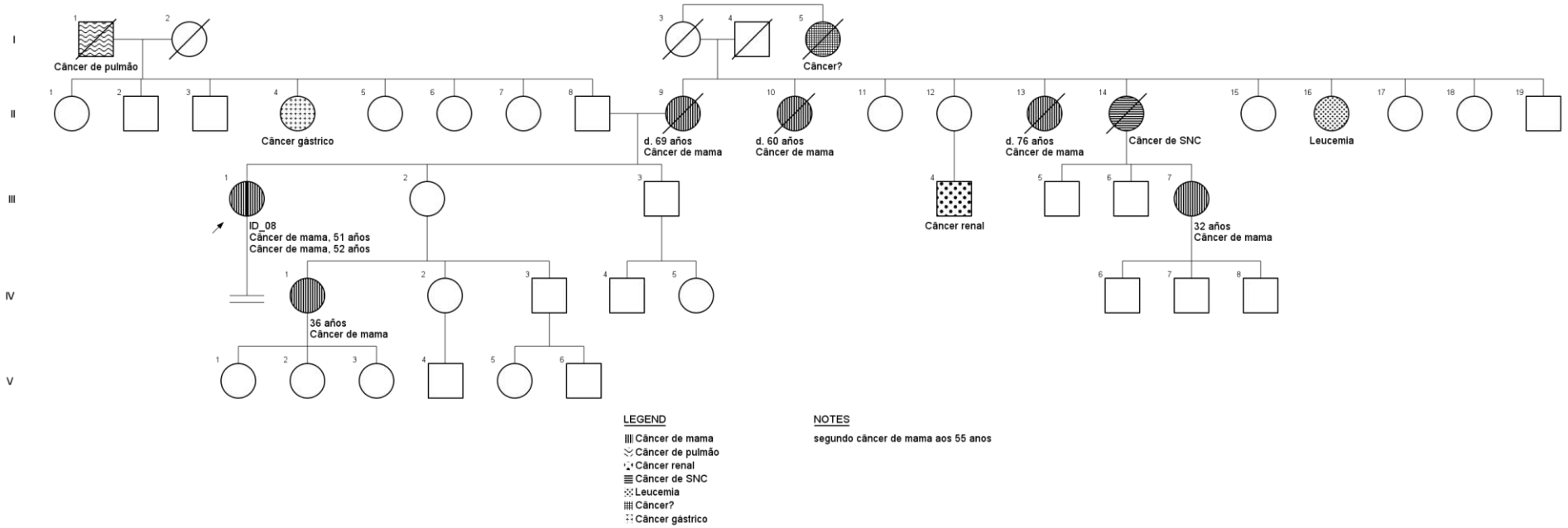
WHO (2016). World Health Organization [Online]. Available: http://www.who.int/en/.

Wong-Brown MW, Meldrum CJ, Carpenter JE, et al (2015). Prevalence of *BRCA1* and *BRCA2* germline mutations in patients with triple-negative breast cancer. *Breast Cancer Res Treat*, **150**, 71-80.

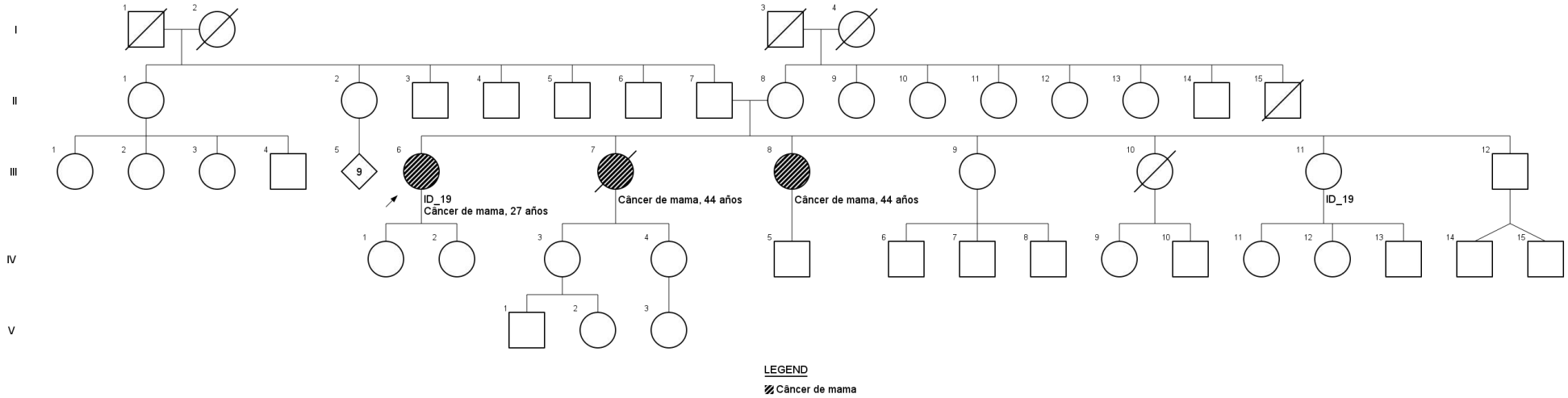
Wooster R, Neuhausen SL, Mangion J, et al (1994). Localization of a breast cancer susceptibility gene, *BRCA2*, to chromosome 13q12-13. *Science*, **265**, 2088-90.

Young SR, Pilarski RT, Donenberg T, et al (2009). The prevalence of *BRCA1* mutations among young women with triple-negative breast cancer. *BMC Cancer*, **9**, 86.

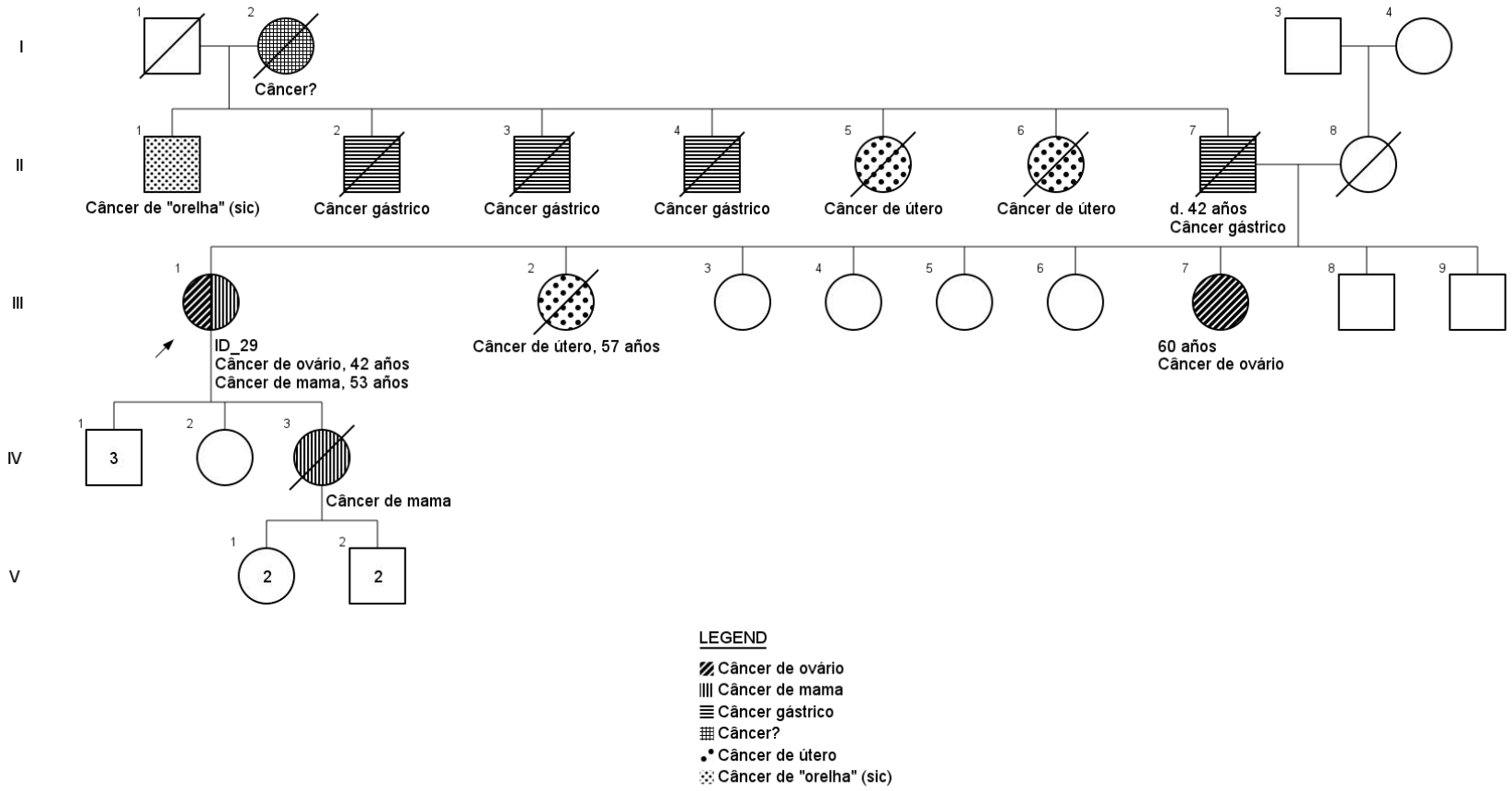
ANEXO VI – Heredograma da paciente ID 8.



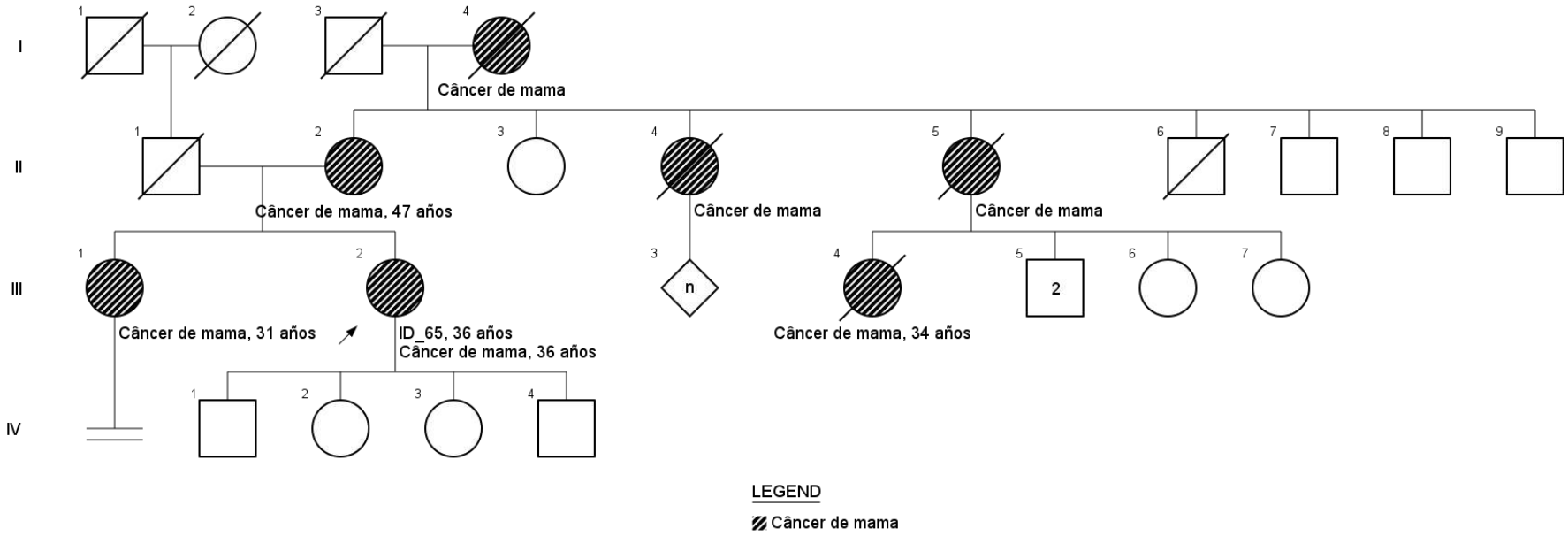
ANEXO VII – Heredograma da paciente ID 19.



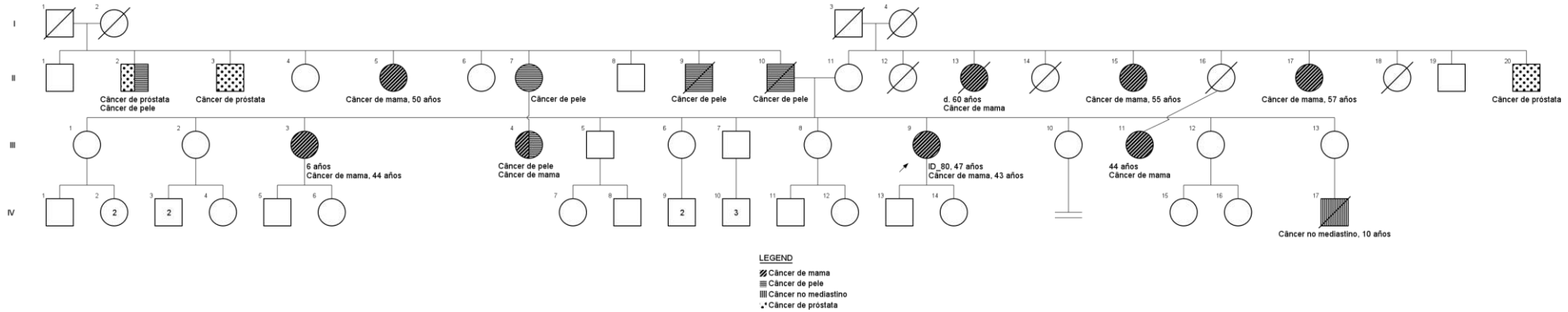
ANEXO VIII – Heredograma da paciente ID 29.



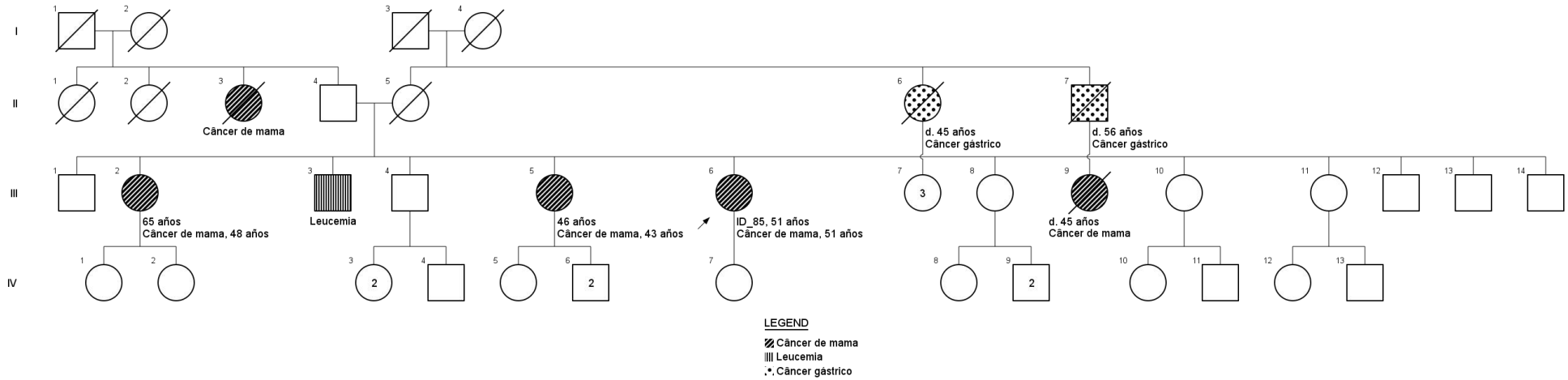
ANEXO IX – Heredograma da paciente ID 65.



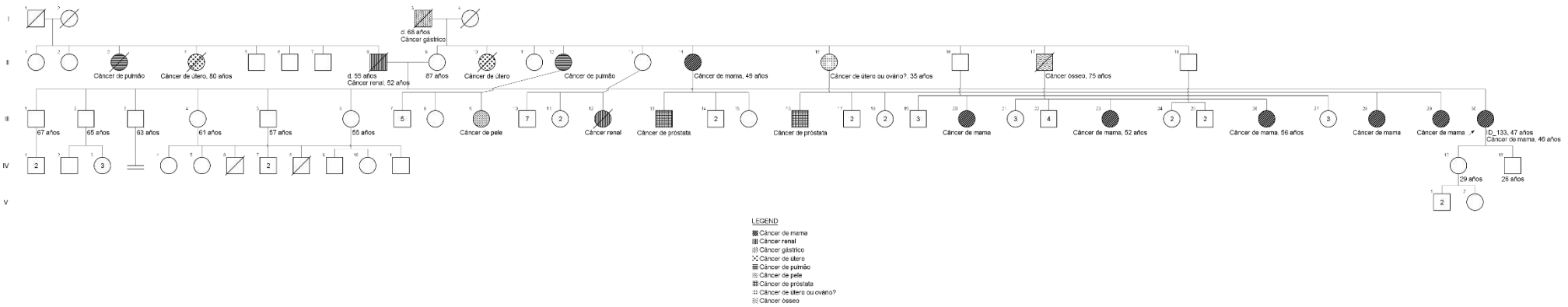
ANEXO X – Heredograma da paciente ID 80.



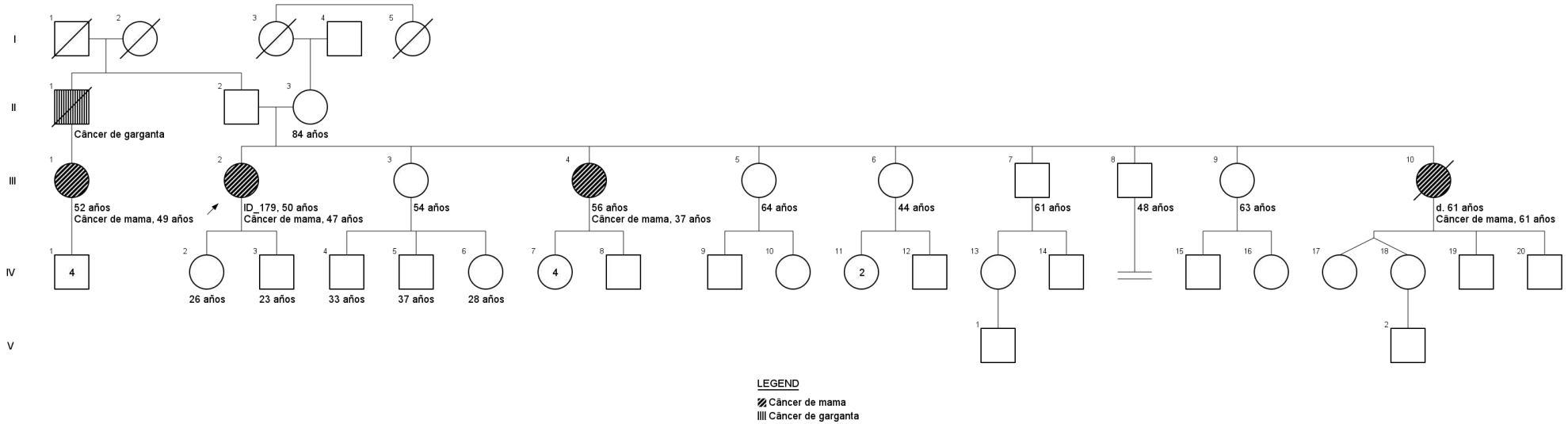
ANEXO XI – Heredograma da paciente ID 85.



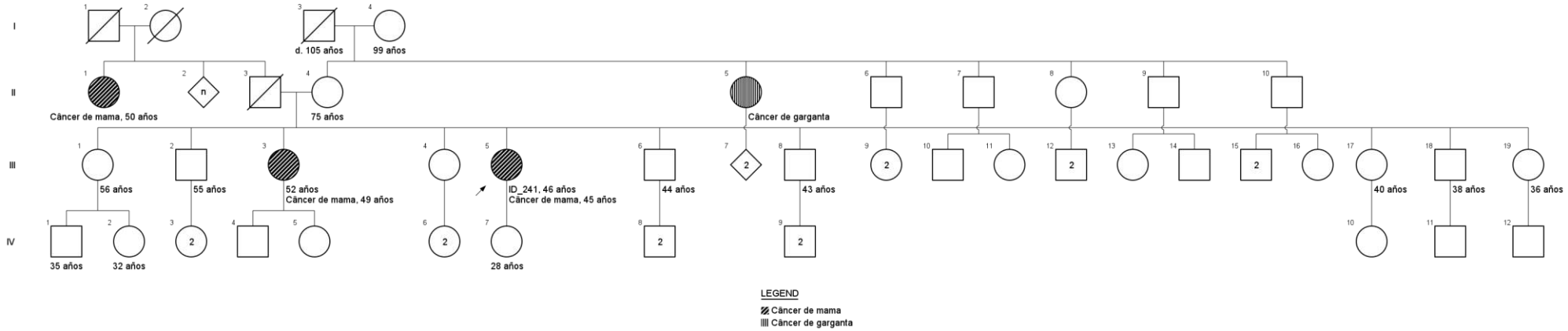
ANEXO XII – Heredograma da paciente ID 133.



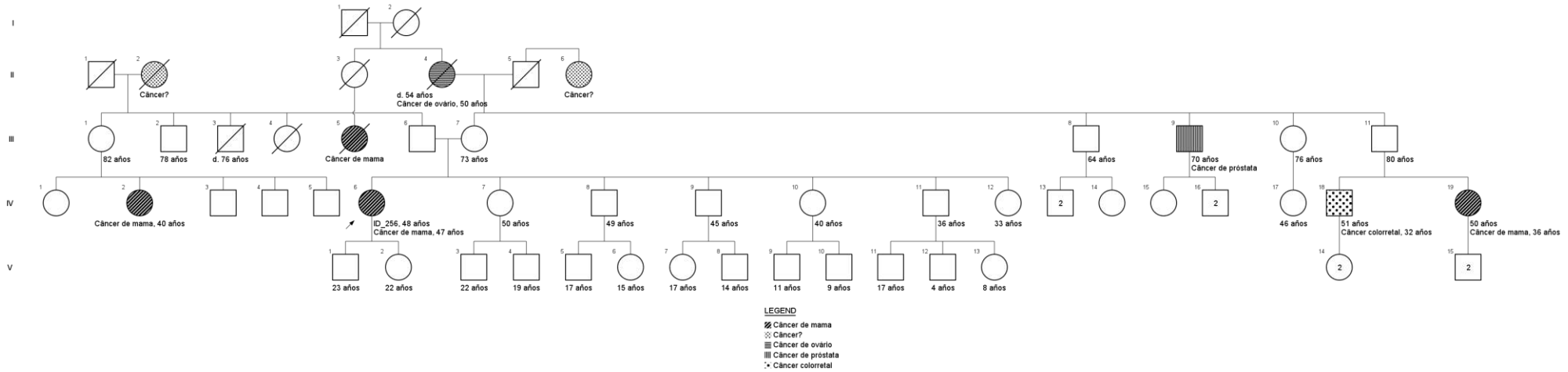
ANEXO XIII – Heredograma da paciente ID 179.



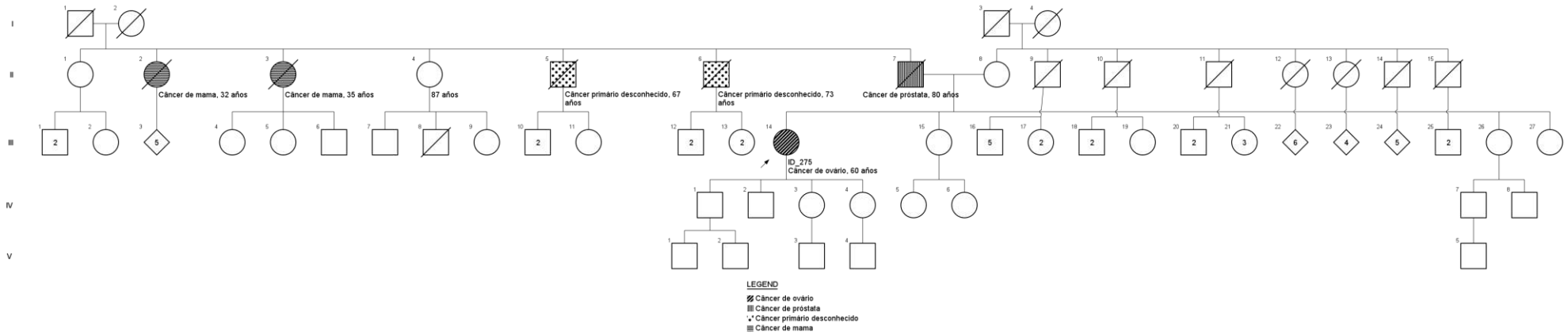
ANEXO XIV – Heredograma da paciente ID 241.



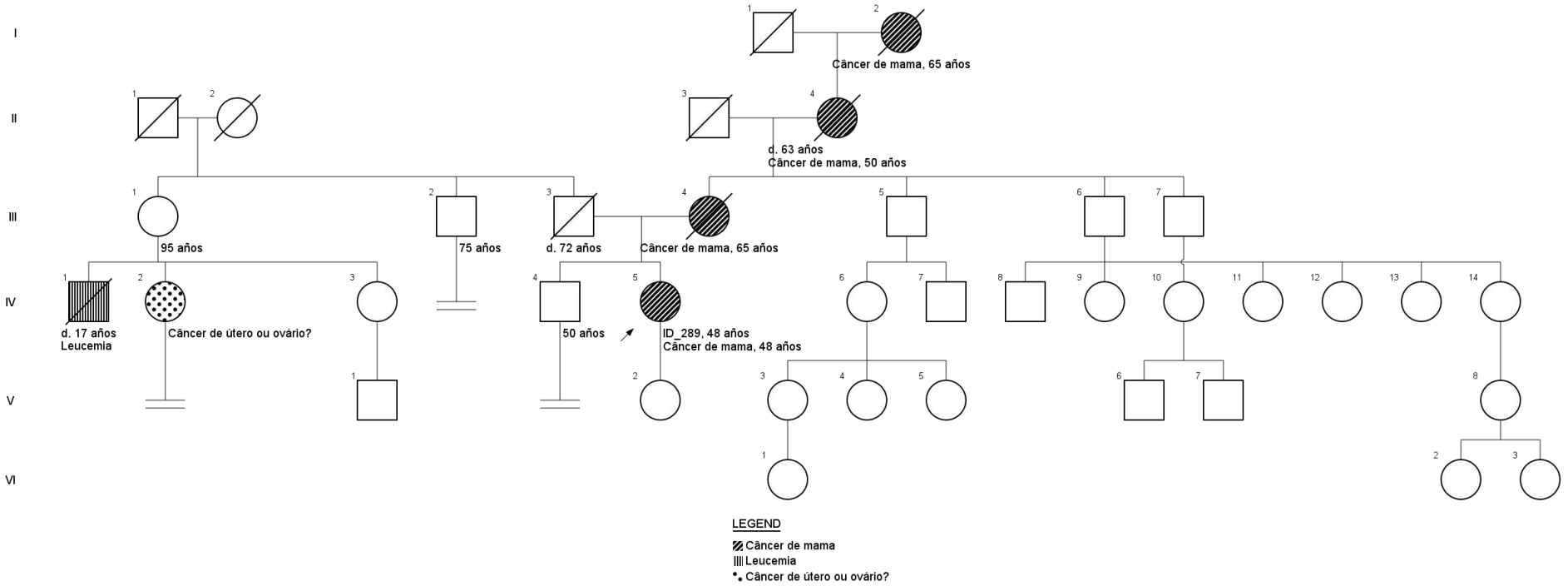
ANEXO XV – Heredograma da paciente ID 256.



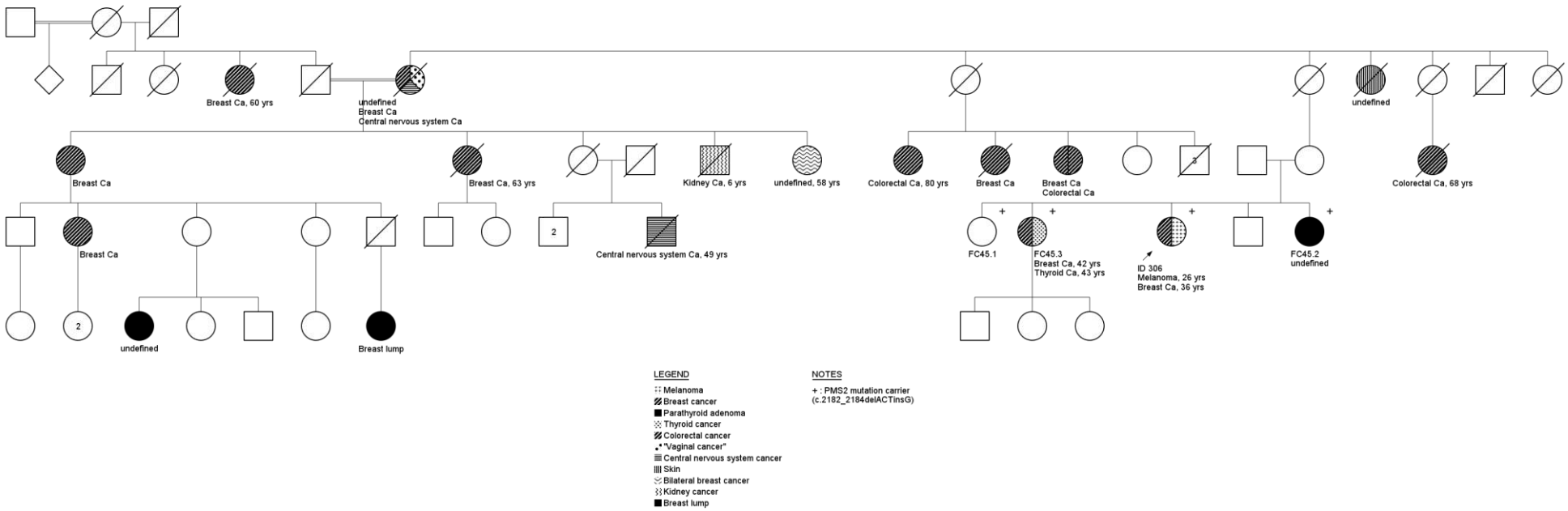
ANEXO XVI – Heredograma da paciente ID 275.



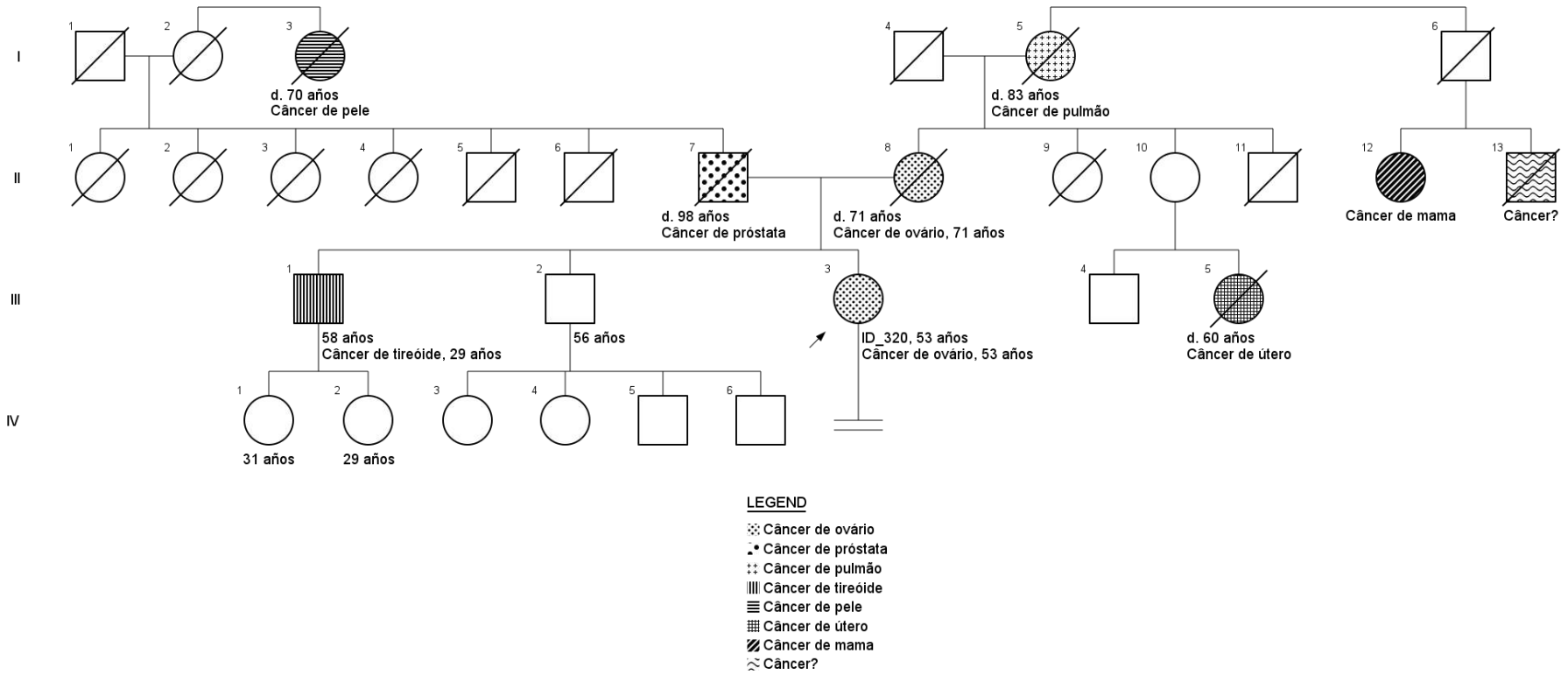
ANEXO XVII – Heredograma da paciente ID 289.



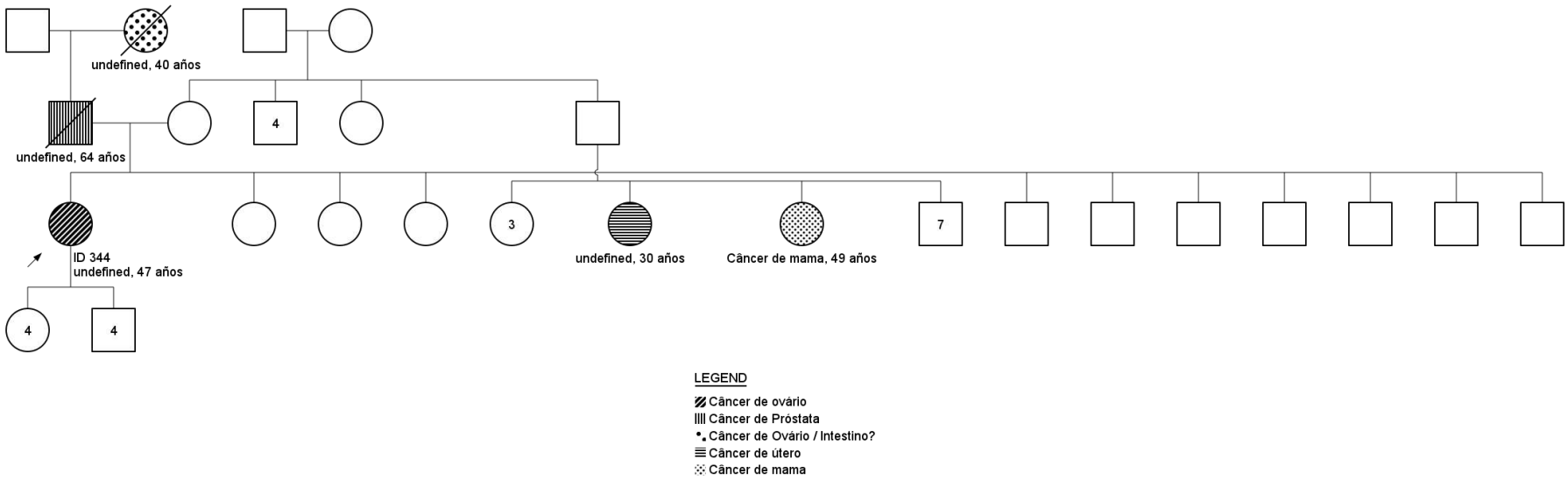
ANEXO XVIII – Heredograma da paciente ID 306.



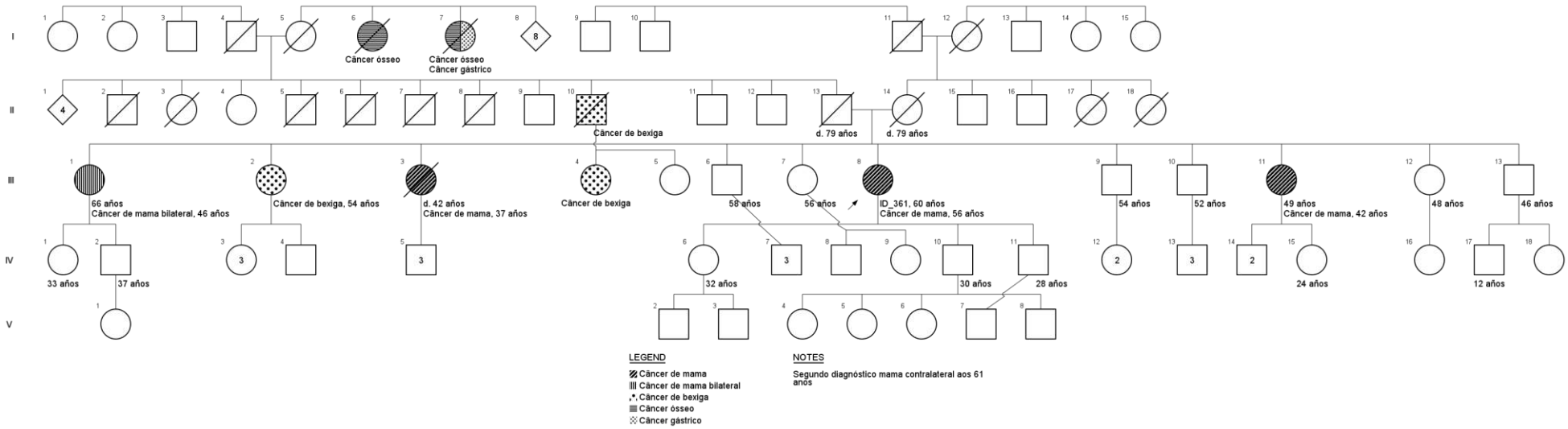
ANEXO XIX – Heredograma da paciente ID 320.



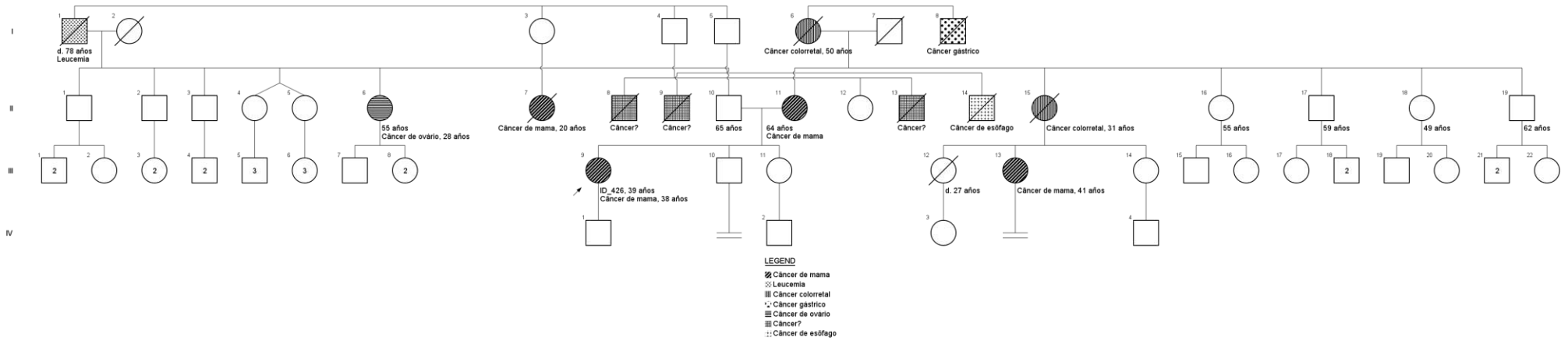
ANEXO XX – Heredograma da paciente ID 344.



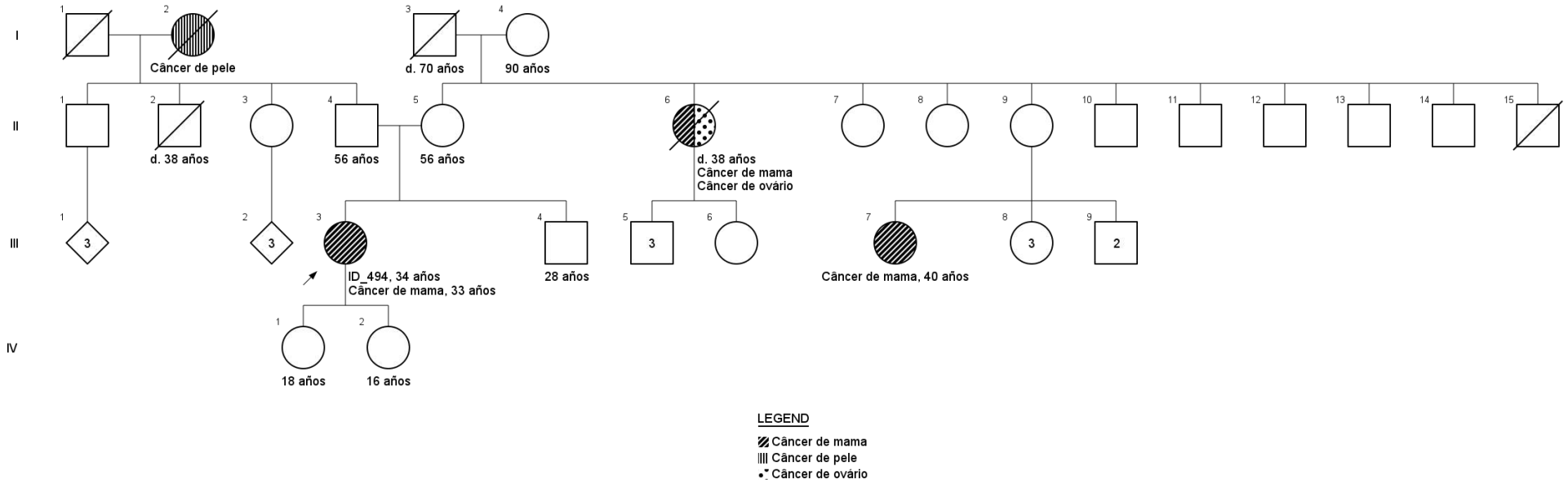
ANEXO XXI – Heredograma da paciente ID 361.



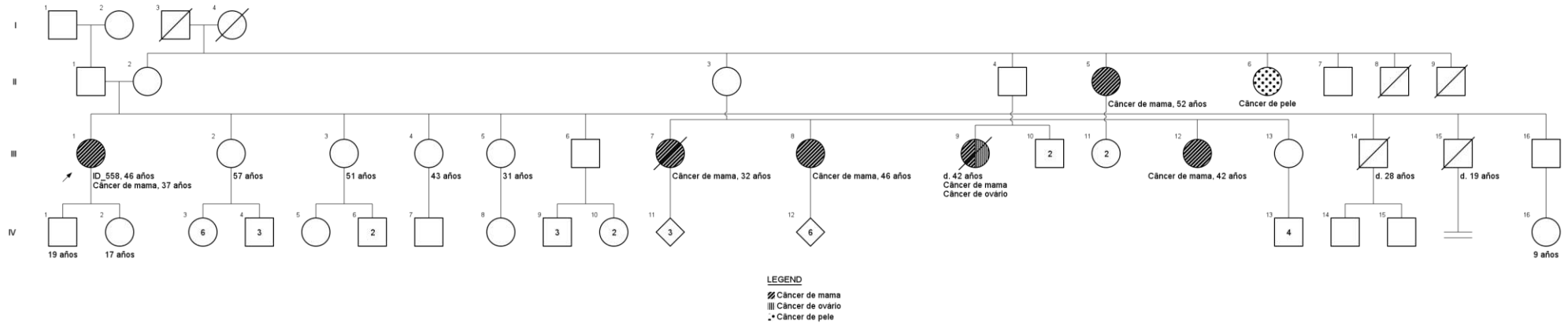
ANEXO XXII – Heredograma da paciente ID 426.



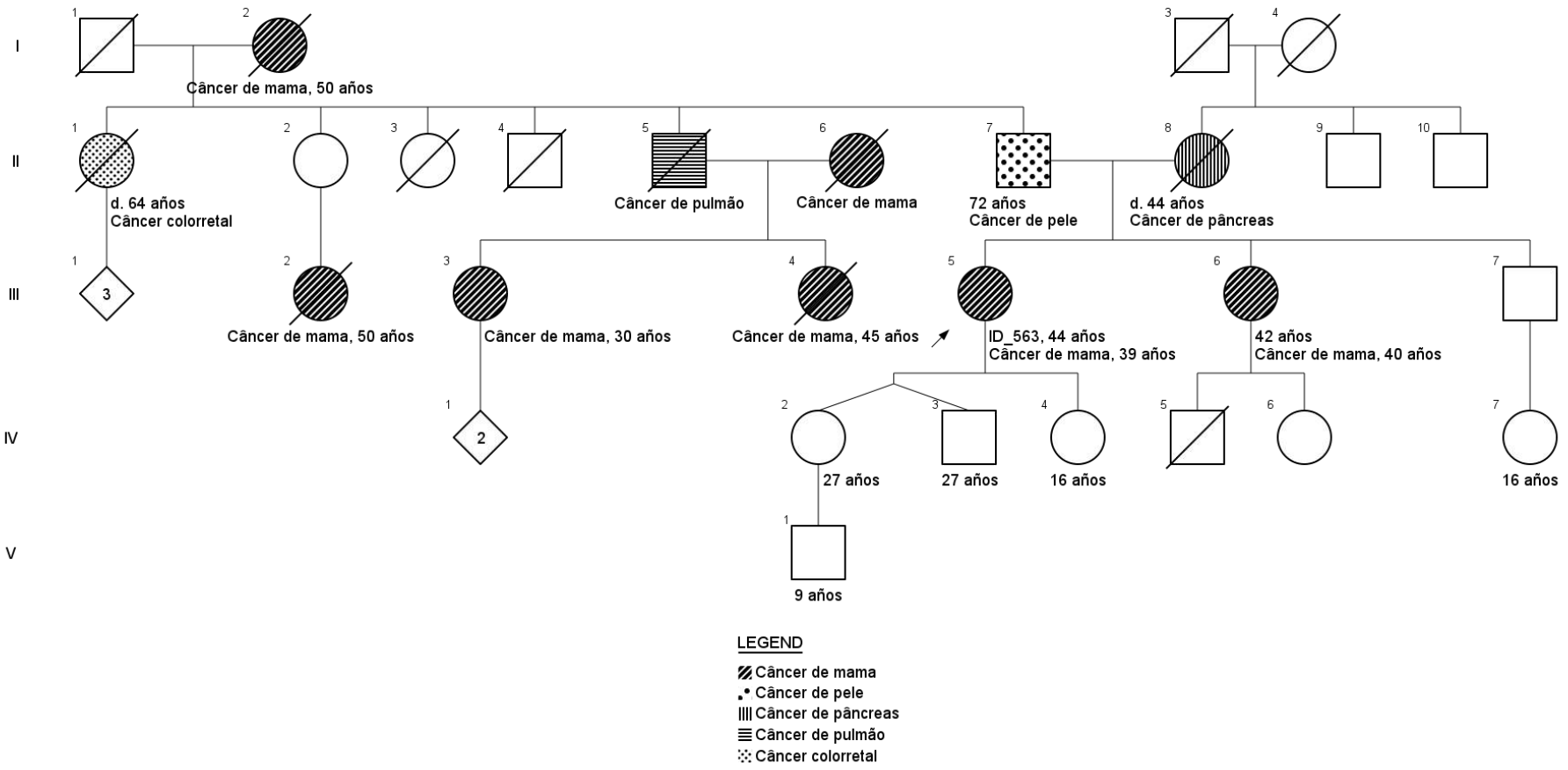
ANEXO XXIII – Heredograma da paciente ID 494.



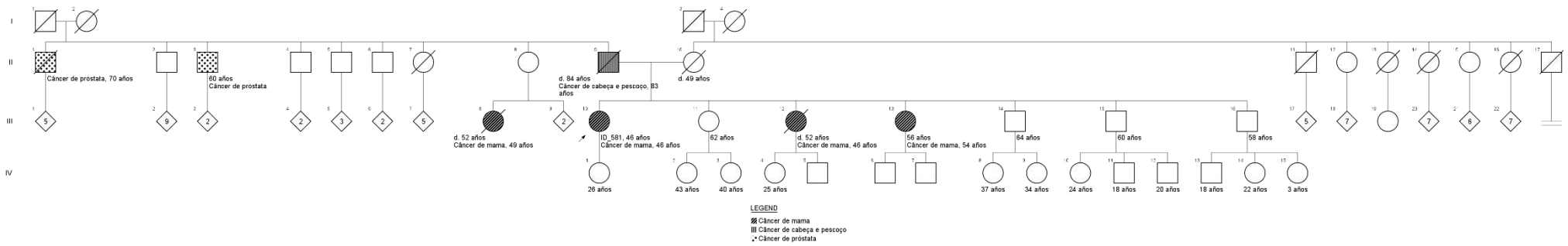
ANEXO XXIV – Heredograma da paciente ID 558.



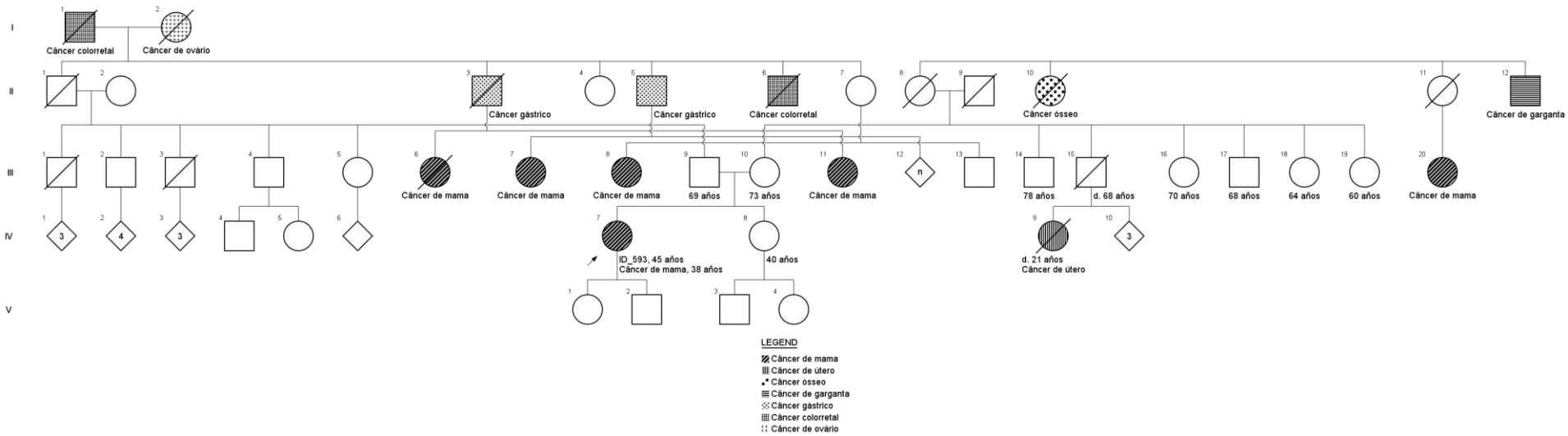
ANEXO XXV – Heredograma da paciente ID 563.



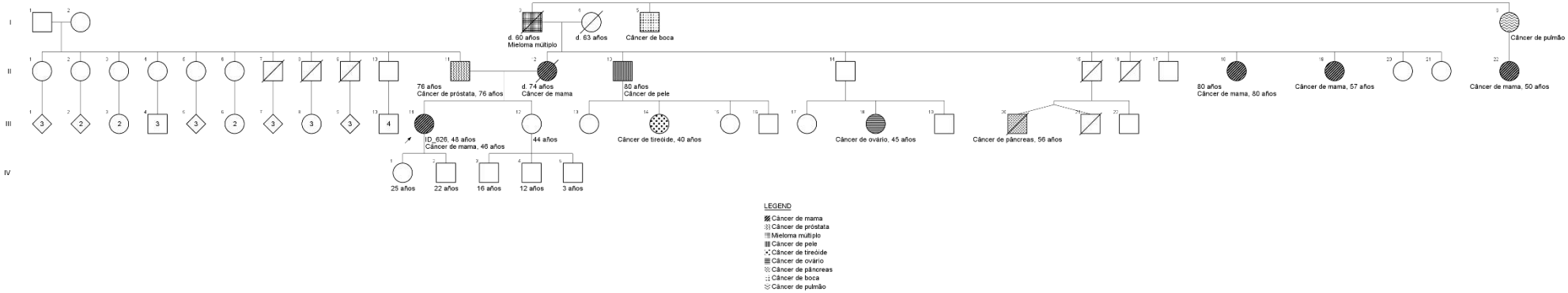
ANEXO XXVI – Heredograma da paciente ID 581.



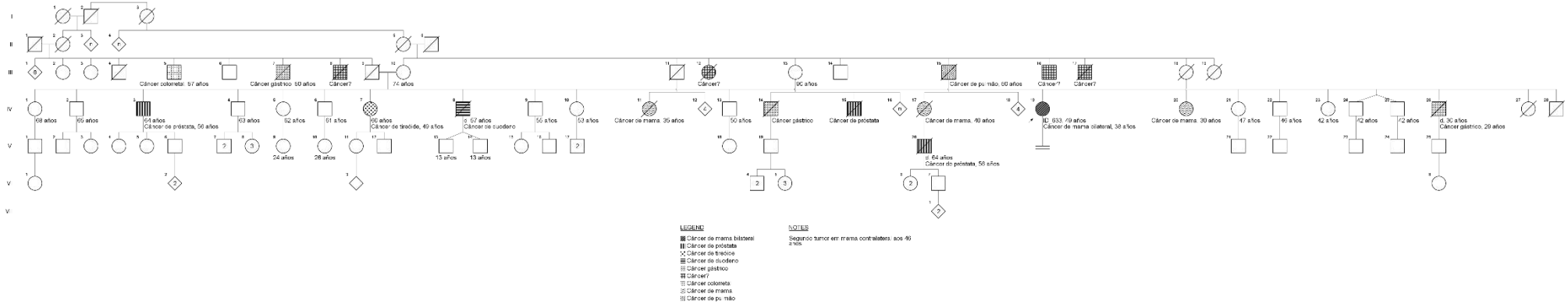
ANEXO XXVII – Heredograma da paciente ID 593.



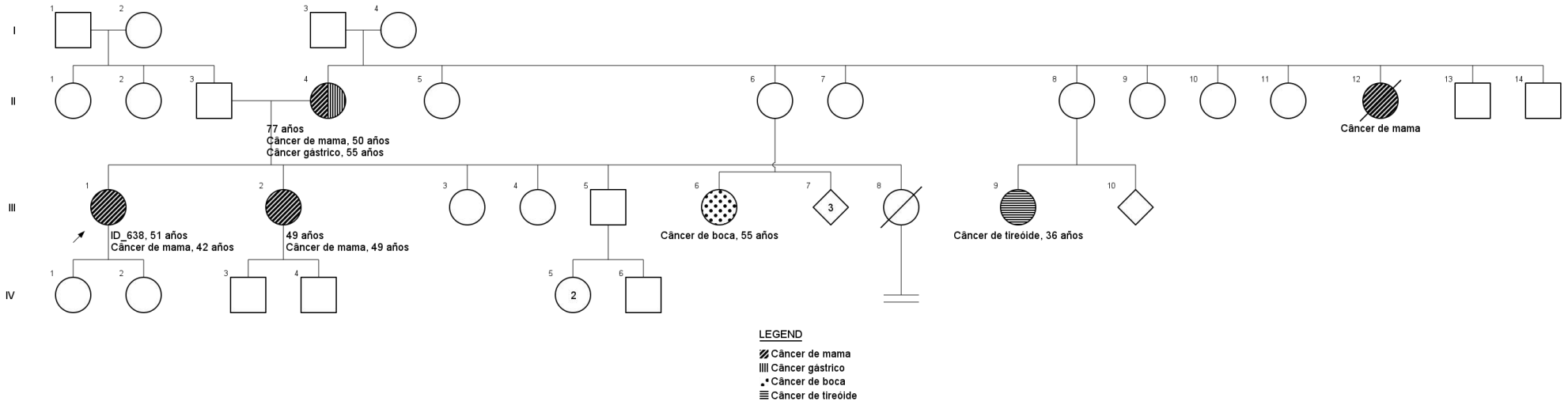
ANEXO XXVIII – Heredograma da paciente ID 626.



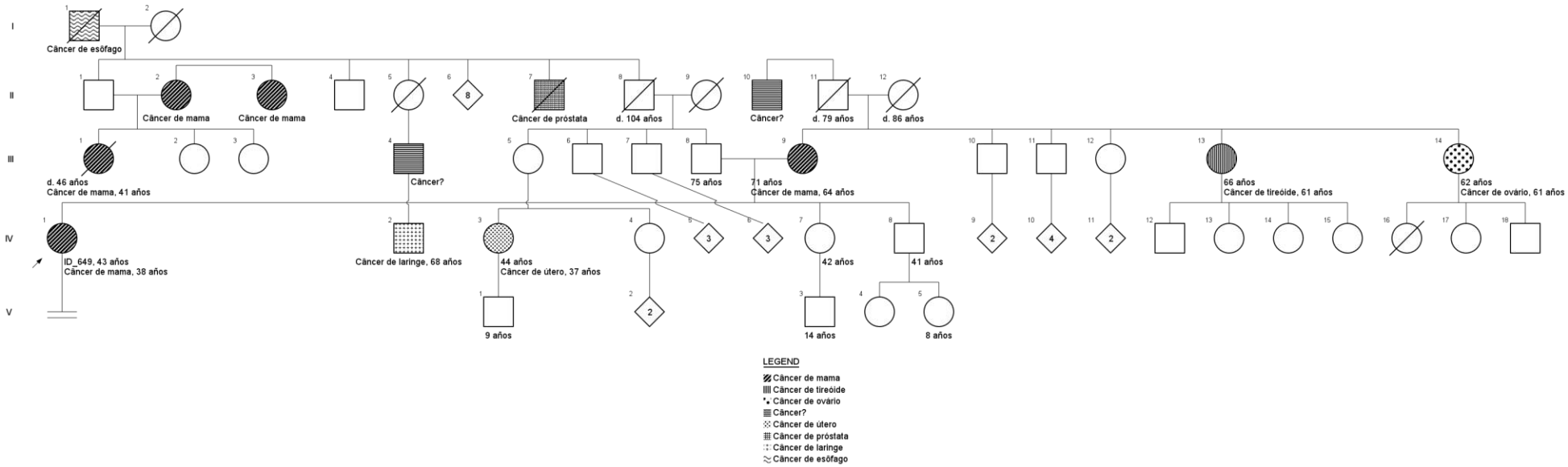
ANEXO XXIX – Heredograma da paciente ID 633.



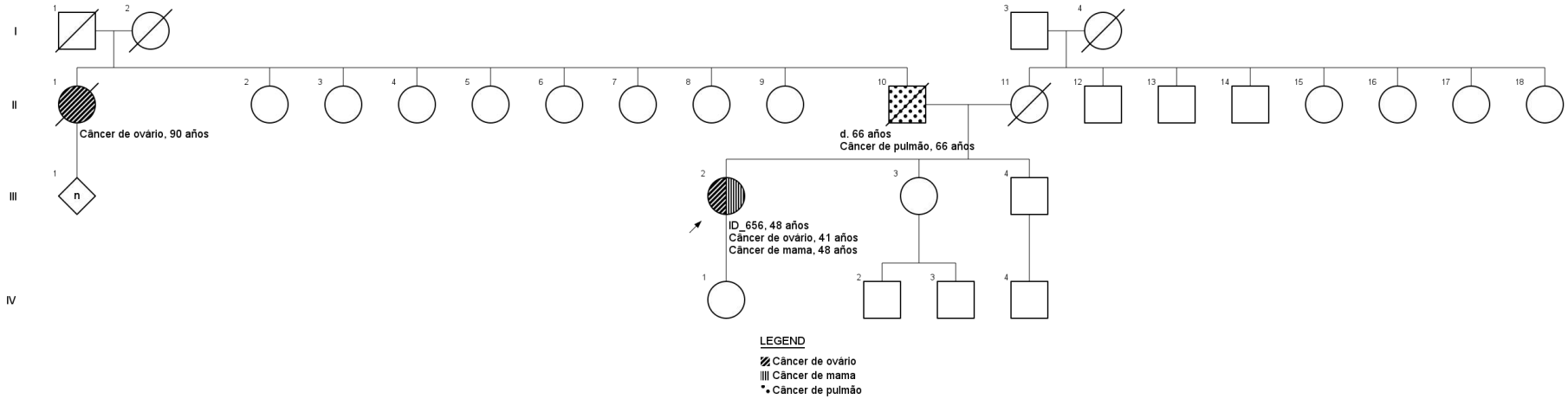
ANEXO XXX – Heredograma da paciente ID 638.



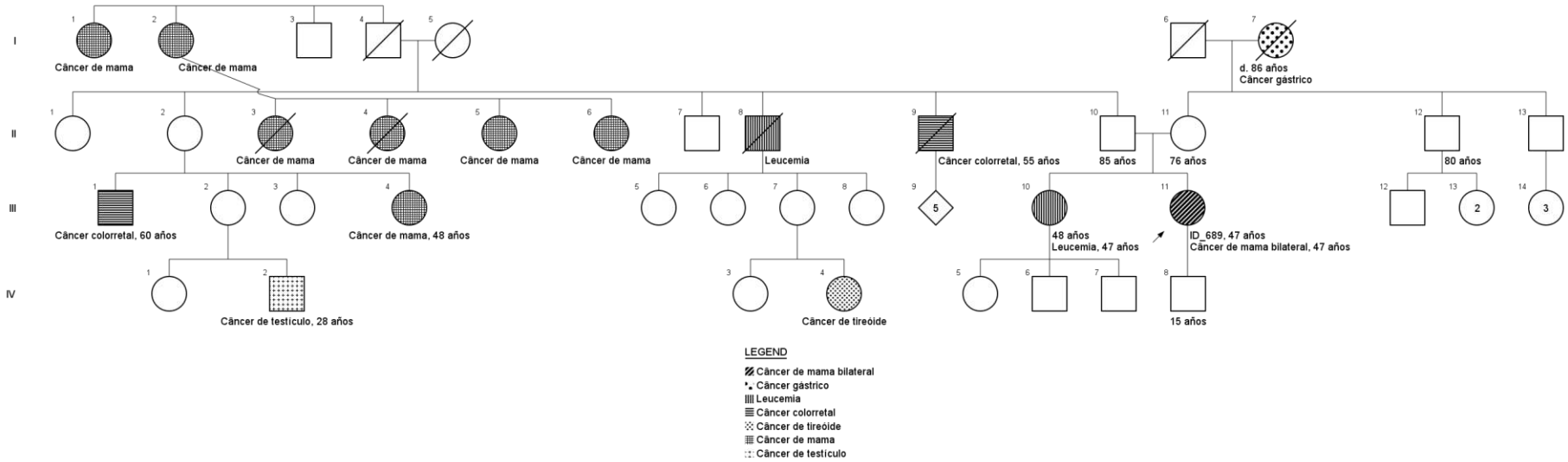
ANEXO XXXI – Heredograma da paciente ID 649.



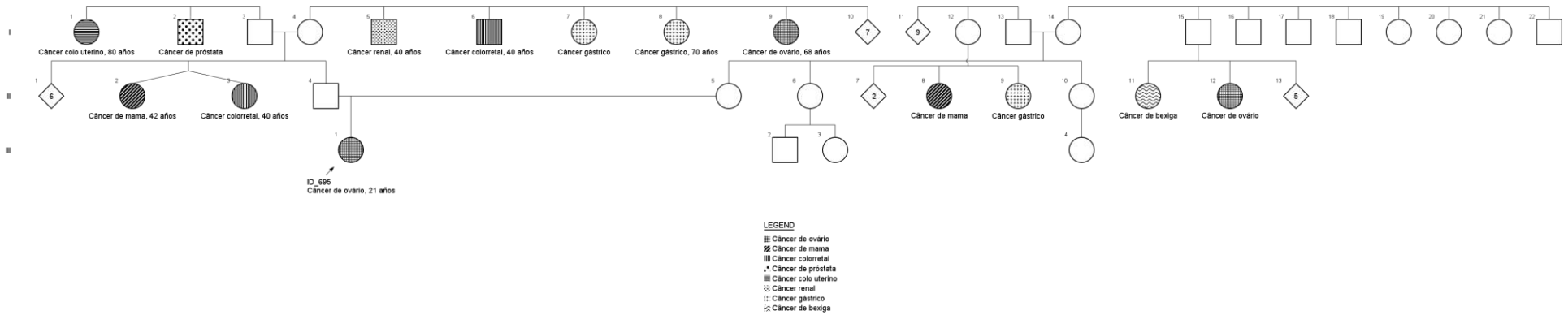
ANEXO XXXII – Heredograma da paciente ID 656.



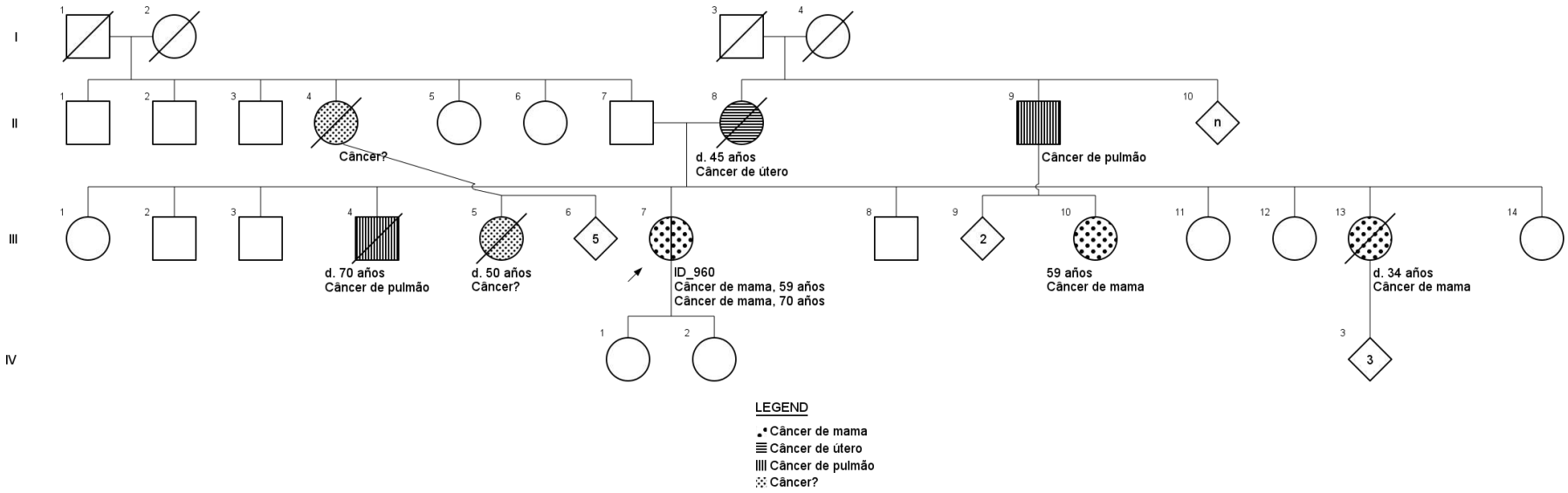
ANEXO XXXIII – Heredograma da paciente ID 689.



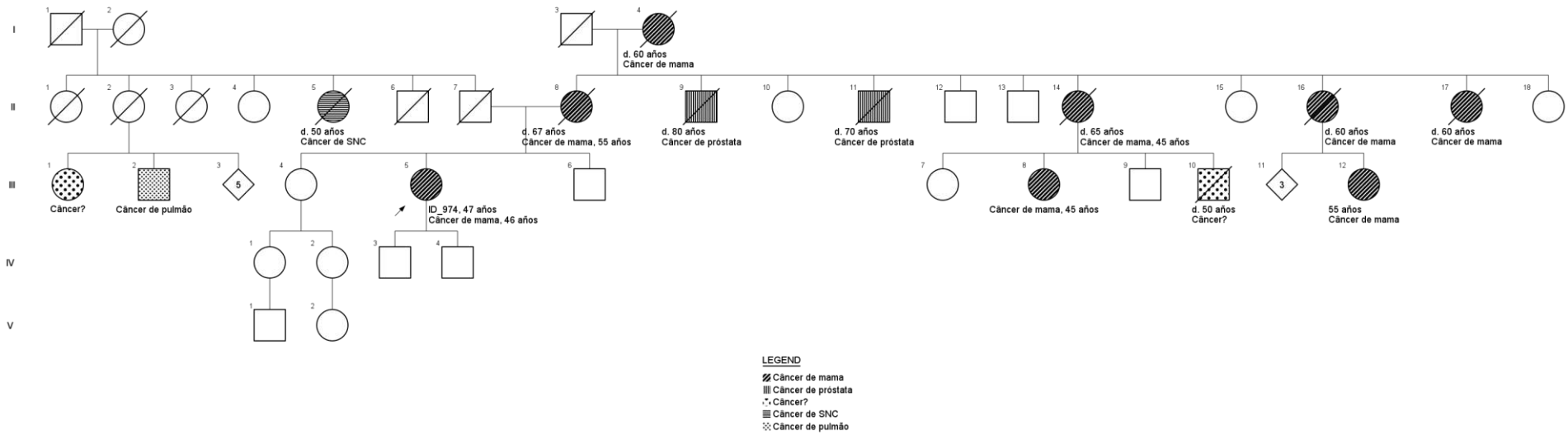
ANEXO XXXIV – Heredograma da paciente ID 695.



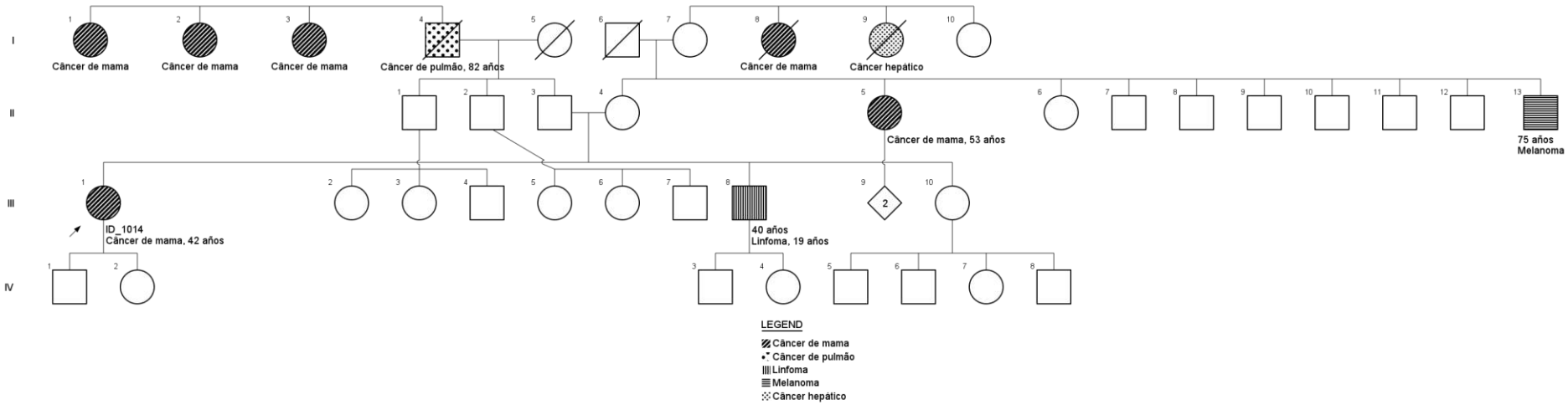
ANEXO XXXV – Heredograma da paciente ID 960.



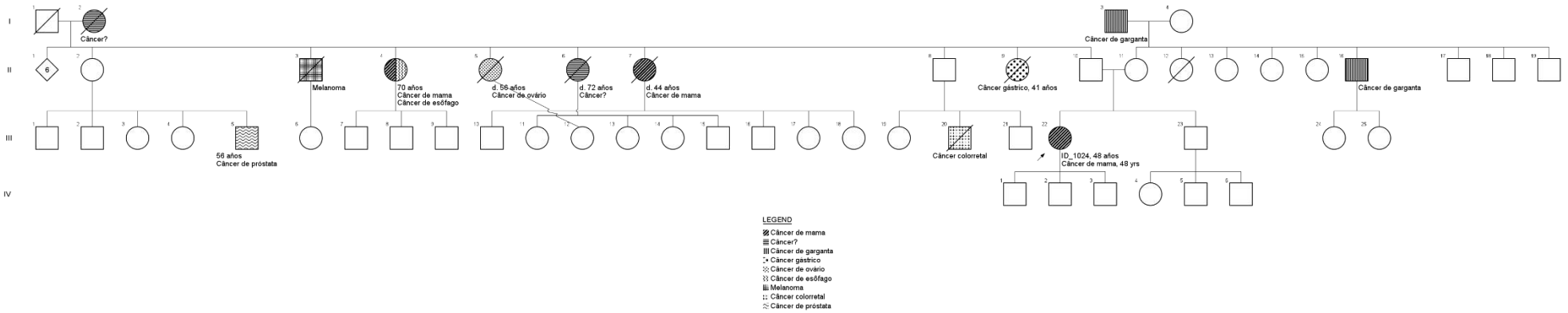
ANEXO XXXVI – Heredograma da paciente ID 974.



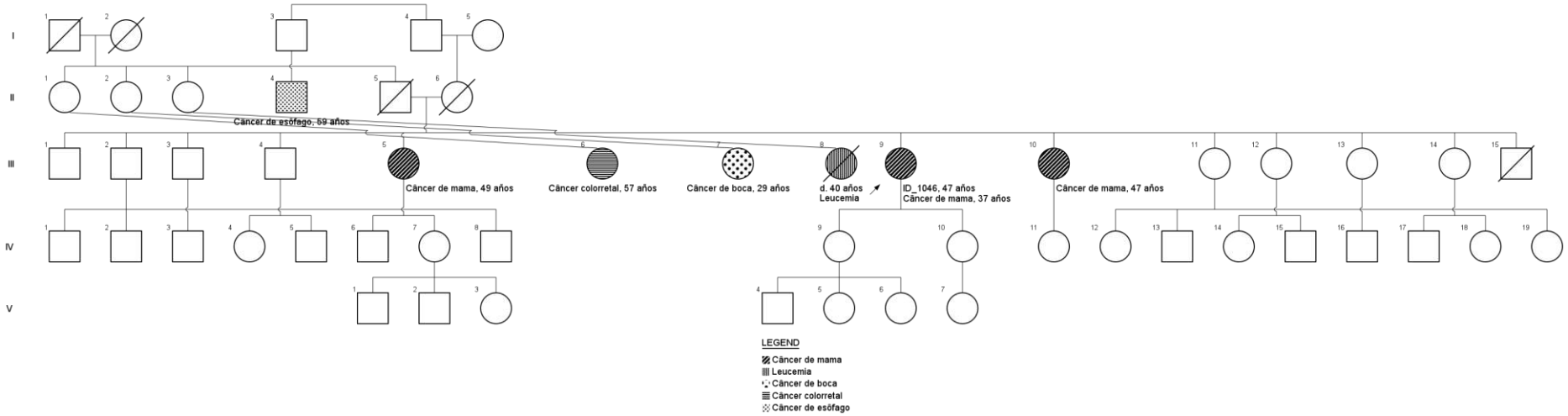
ANEXO XXXVII – Heredograma da paciente ID 1014.



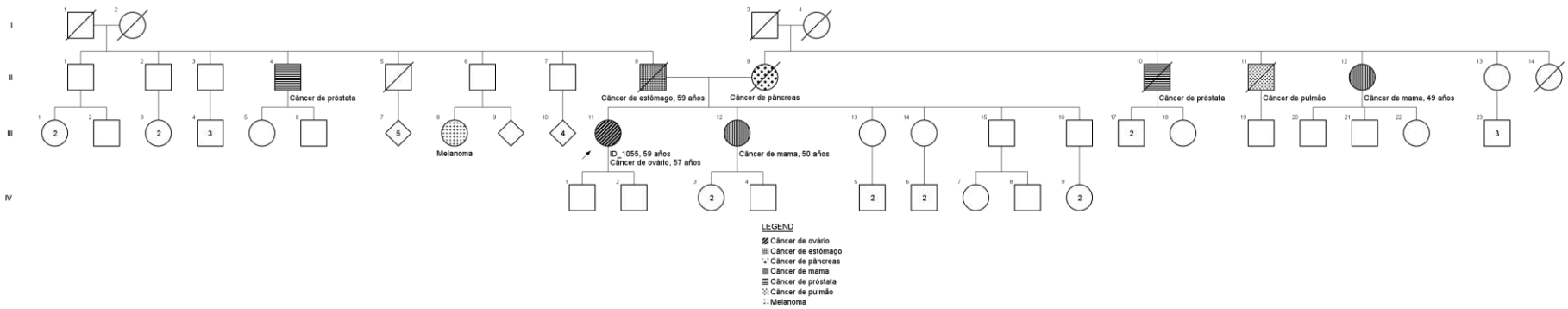
ANEXO XXXVIII – Heredograma da paciente ID 1024.



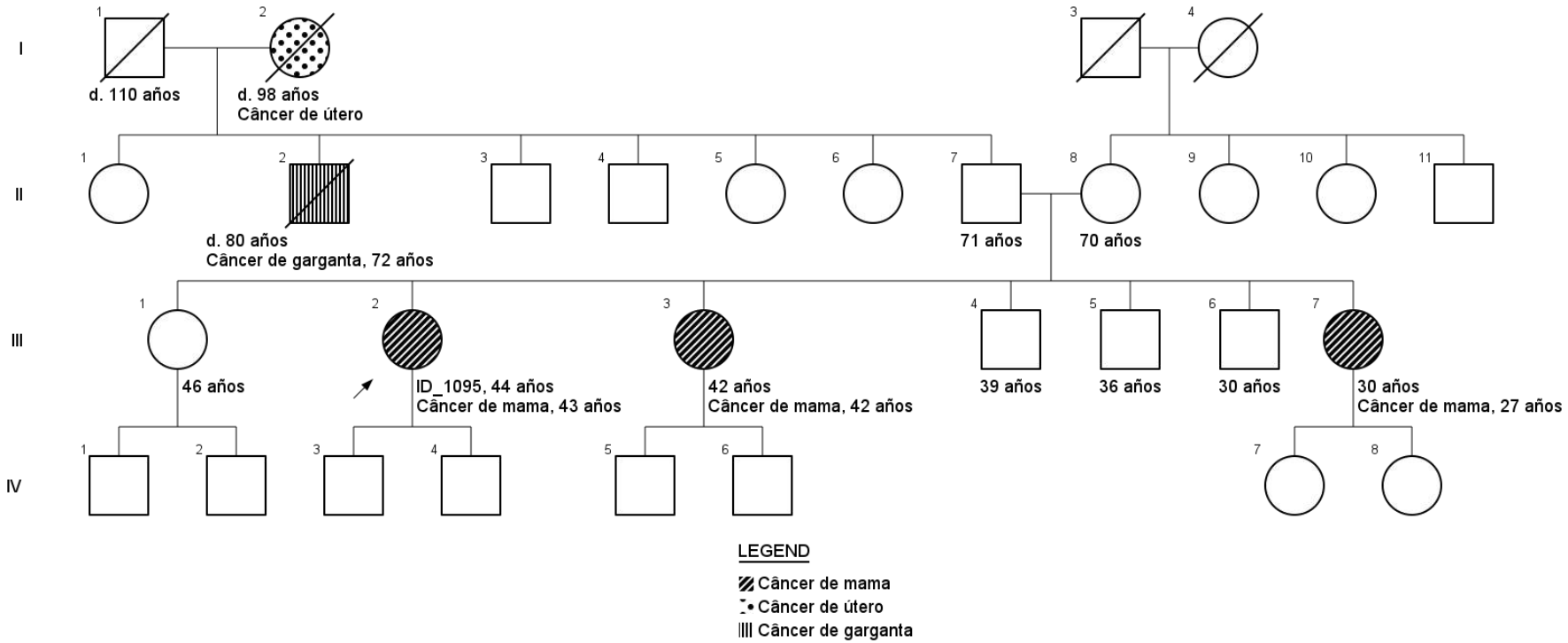
ANEXO XXXIX – Heredograma da paciente ID 1046.



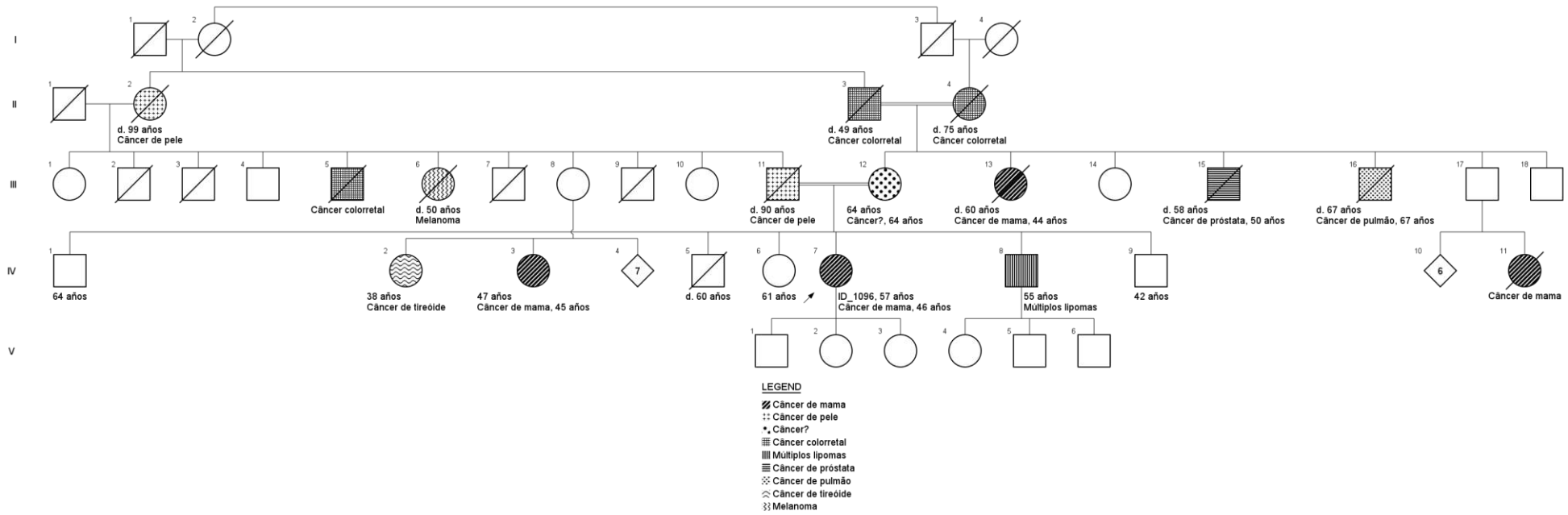
ANEXO XL – Heredograma da paciente ID 1055.



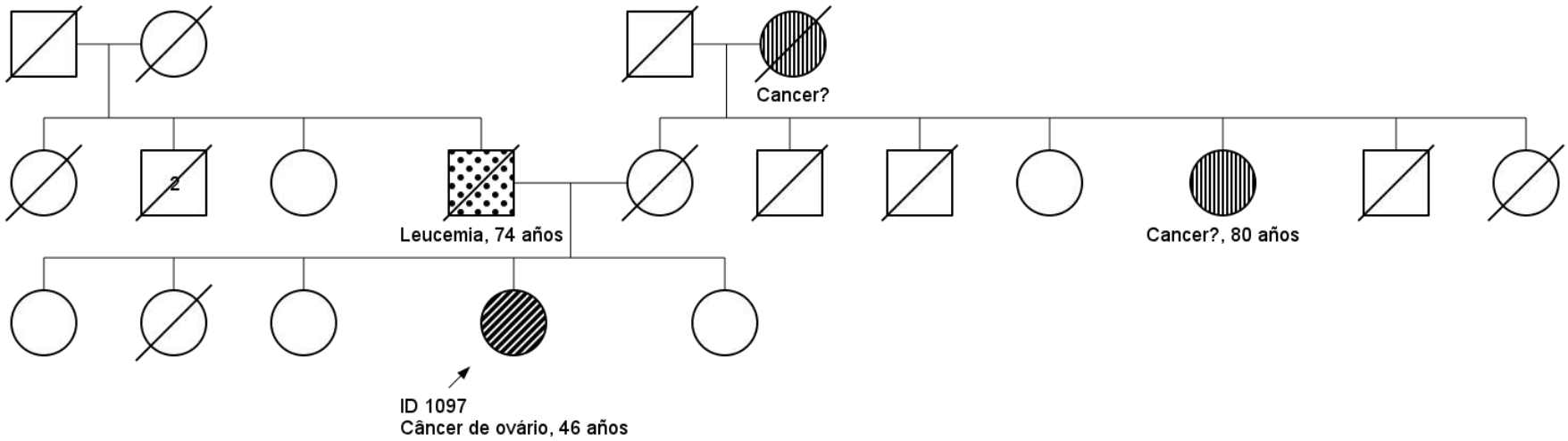
ANEXO XLI – Heredograma da paciente ID 1095.



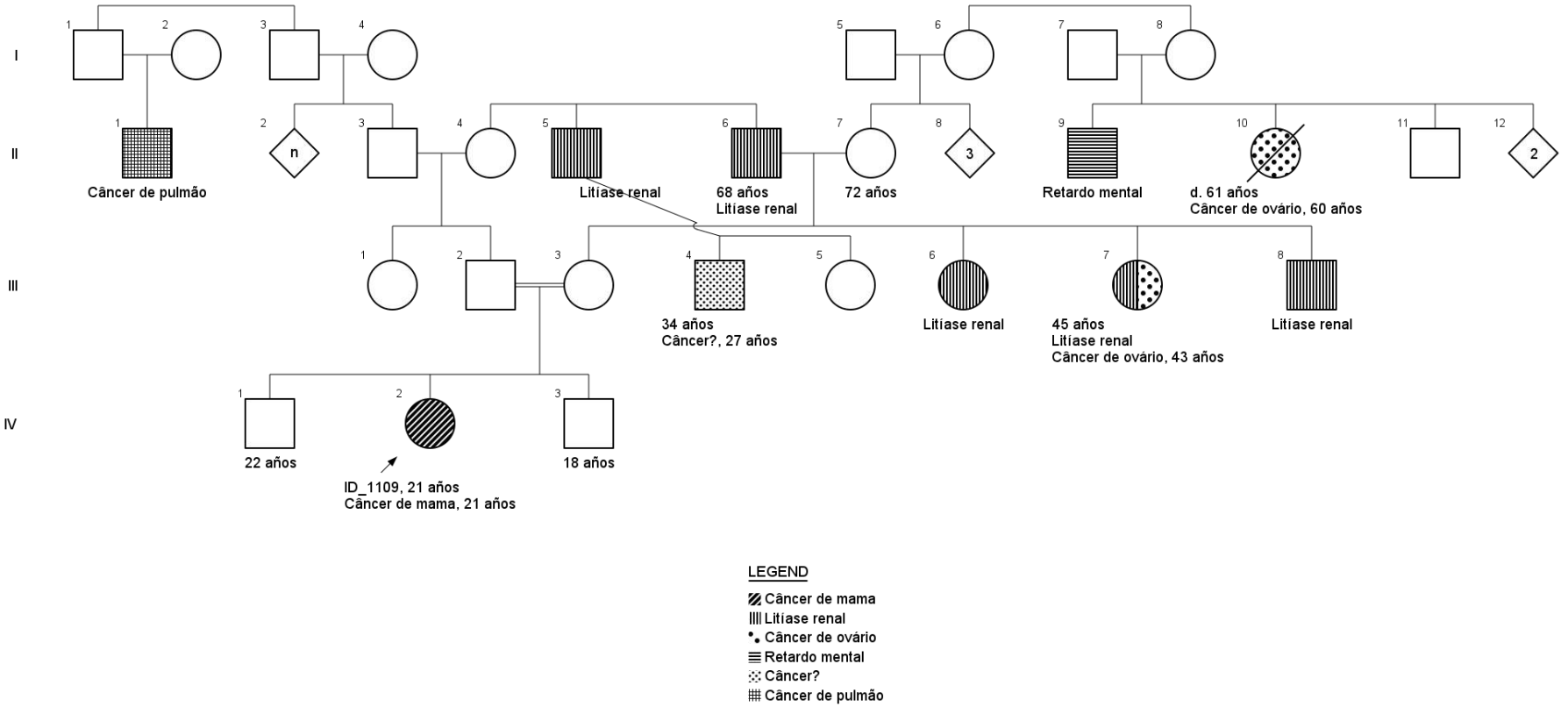
ANEXO XLII – Heredograma da paciente ID 1096.



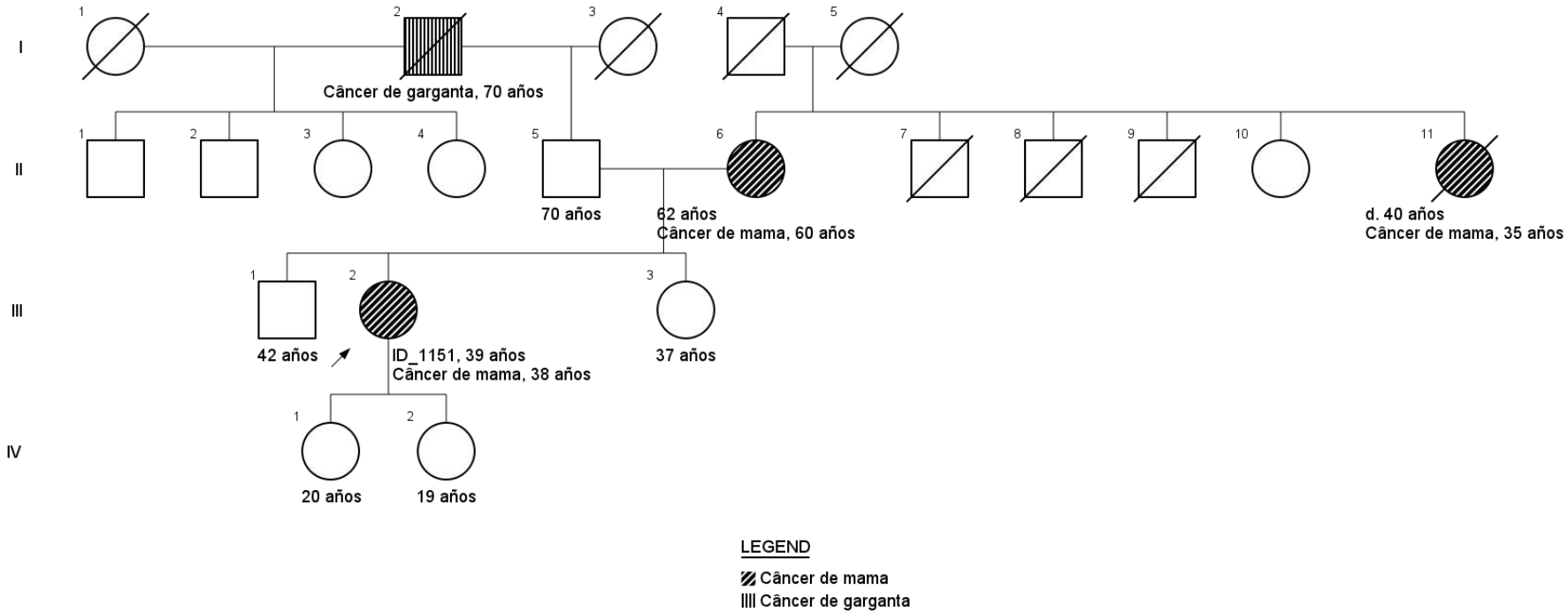
ANEXO XLIII – Heredograma da paciente ID 1097.



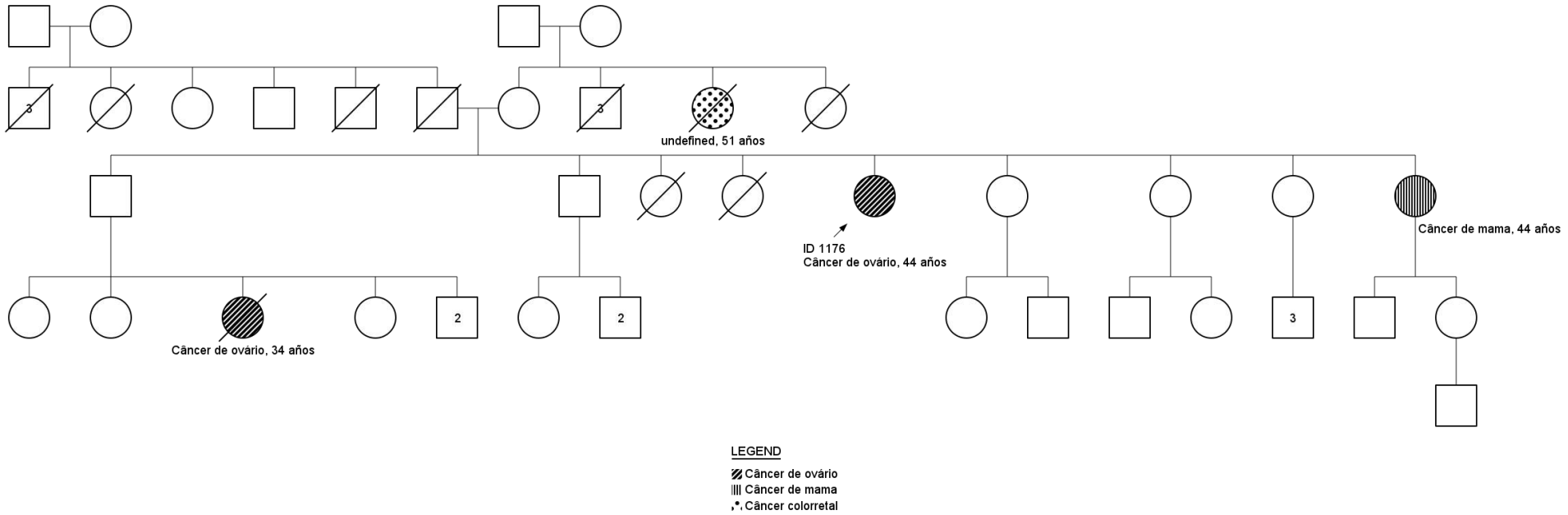
ANEXO XLIV – Heredograma da paciente ID 1109.



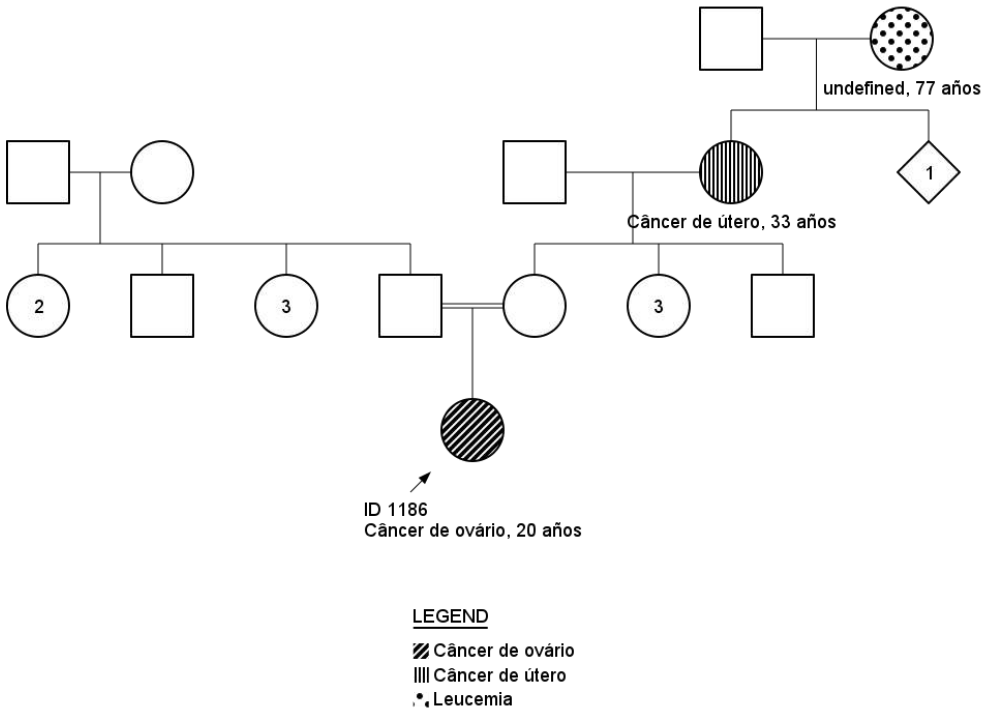
ANEXO XLV – Heredograma da paciente ID 1151.



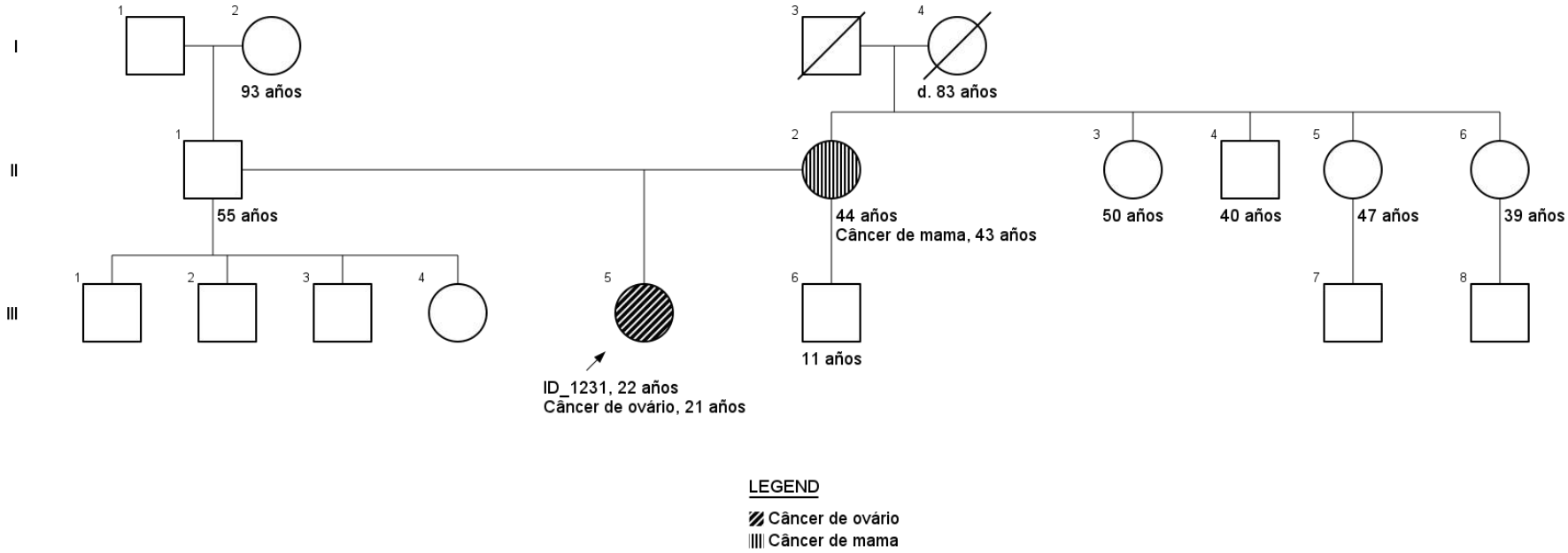
ANEXO XLVI – Heredograma da paciente ID 1176.



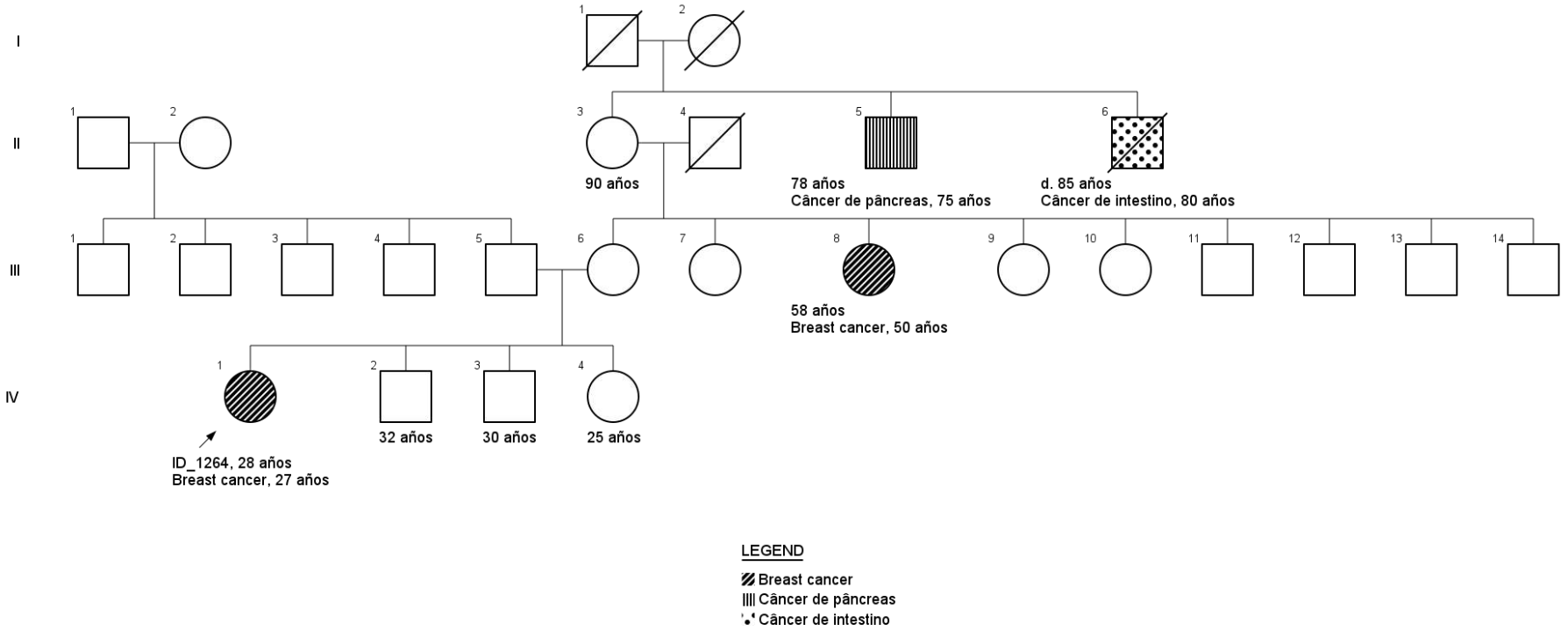
ANEXO XLVII – Heredograma da paciente ID 1186.



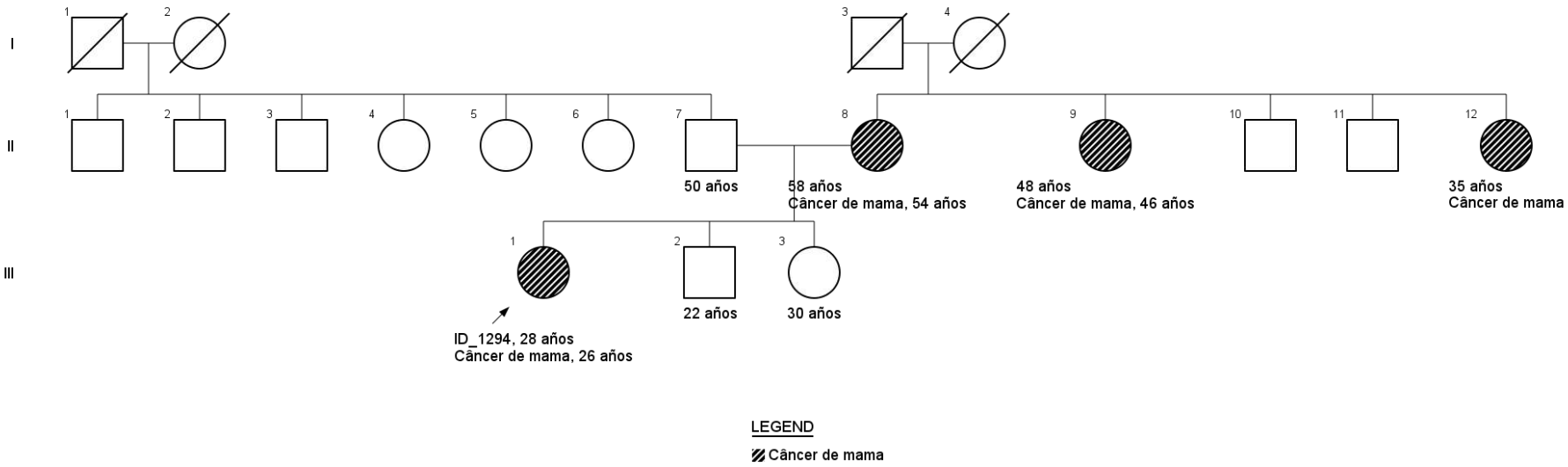
ANEXO XLVIII – Heredograma da paciente ID 1231.



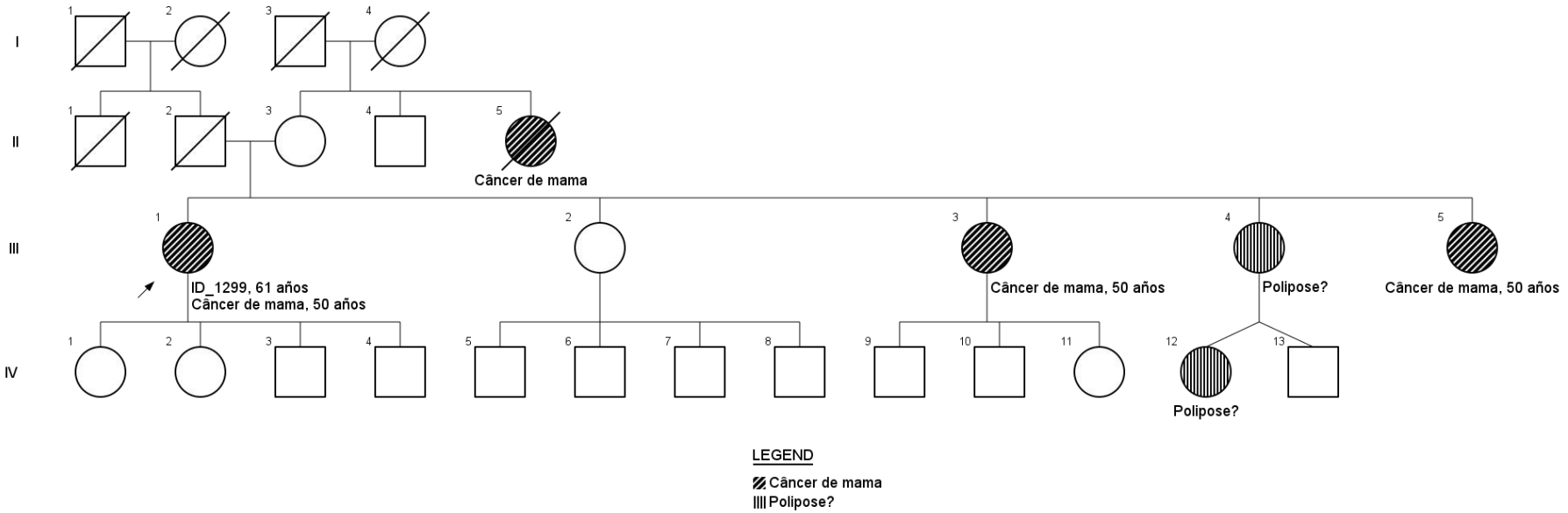
ANEXO XLIX – Heredograma da paciente ID 1264.



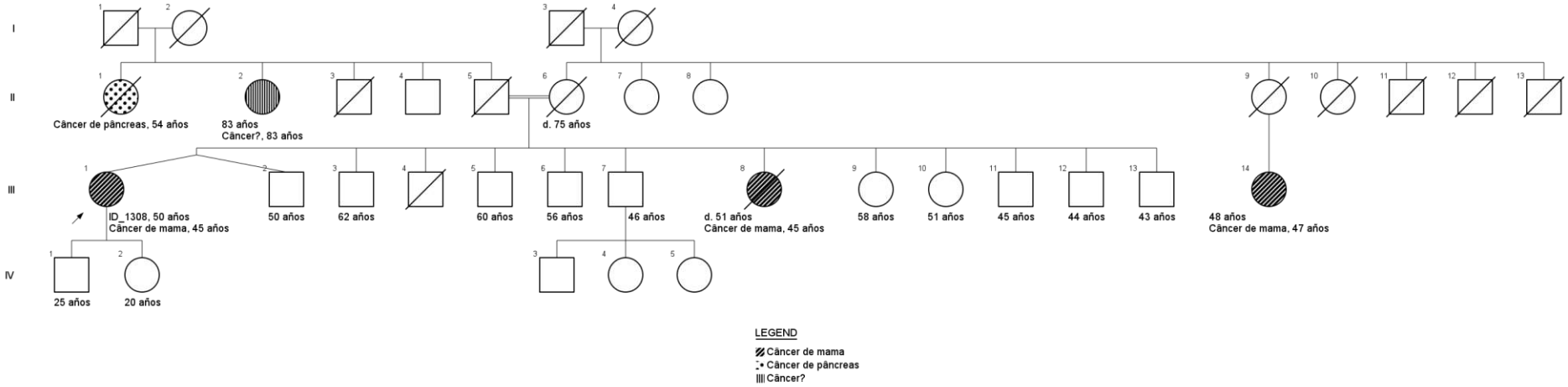
ANEXO L – Heredograma da paciente ID 1294.



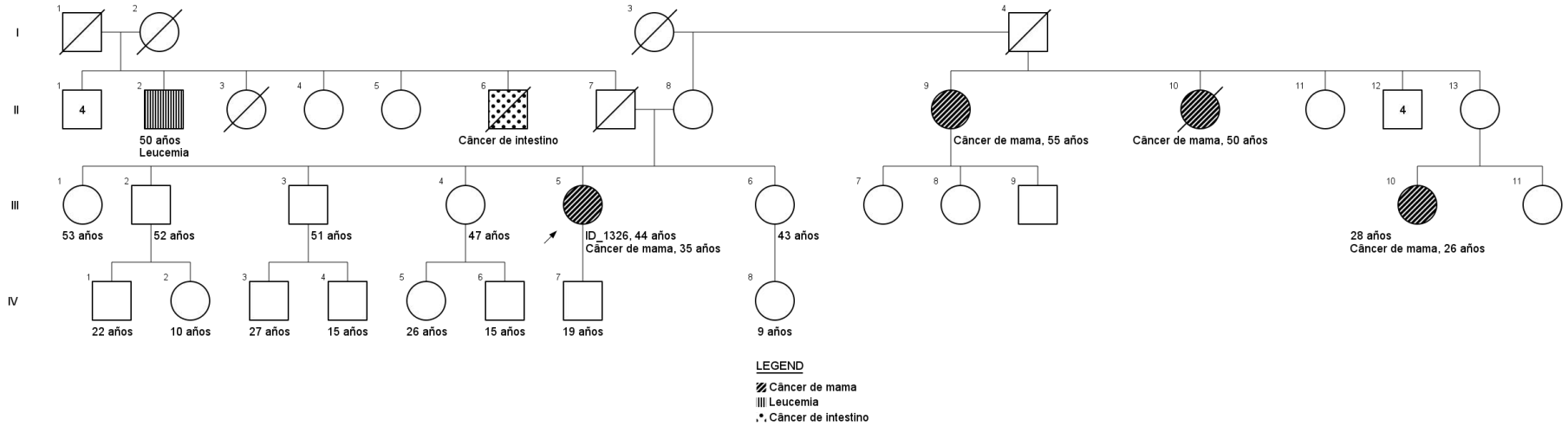
ANEXO LI – Heredograma da paciente ID 1299.



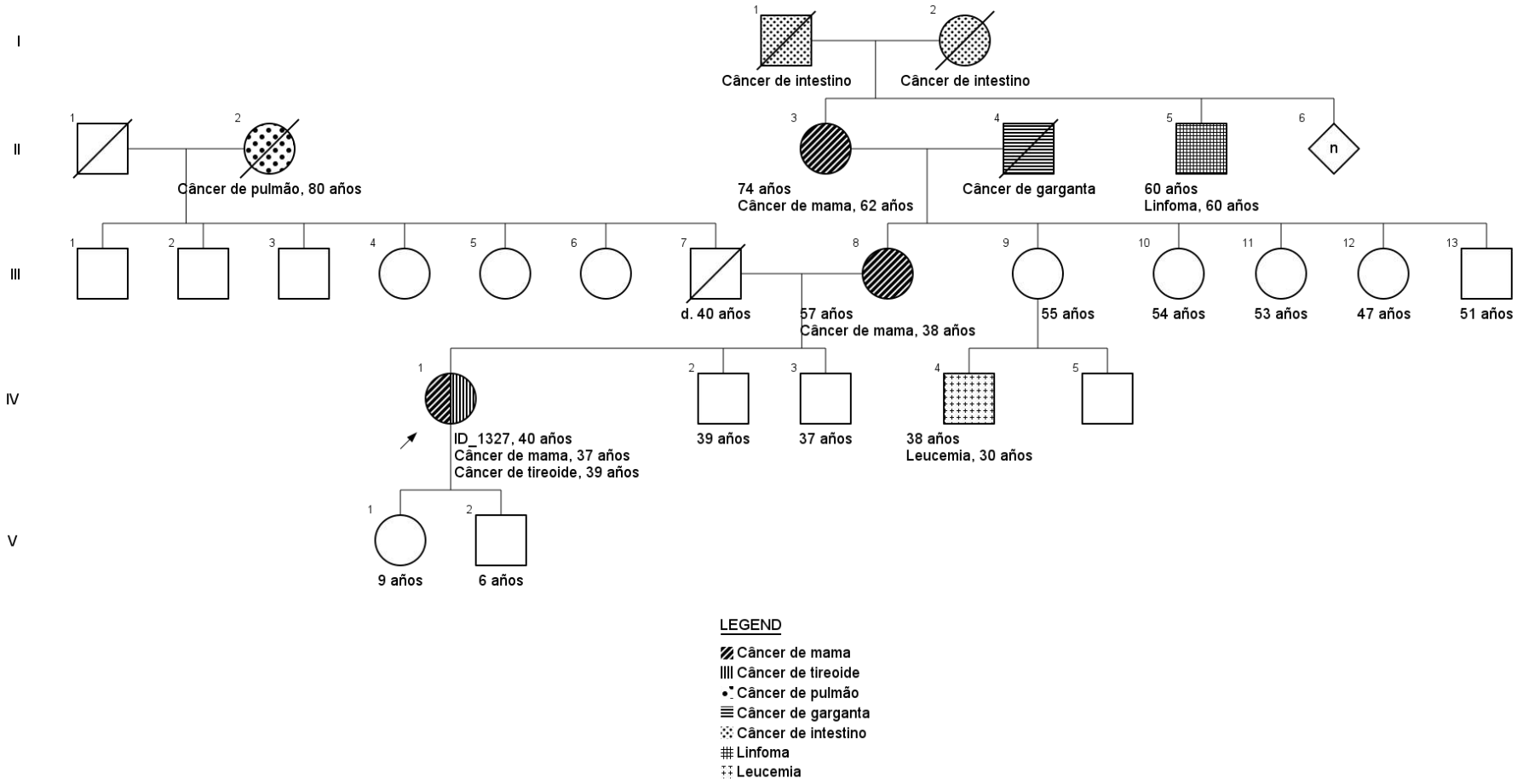
ANEXO LII – Heredograma da paciente ID 1308.



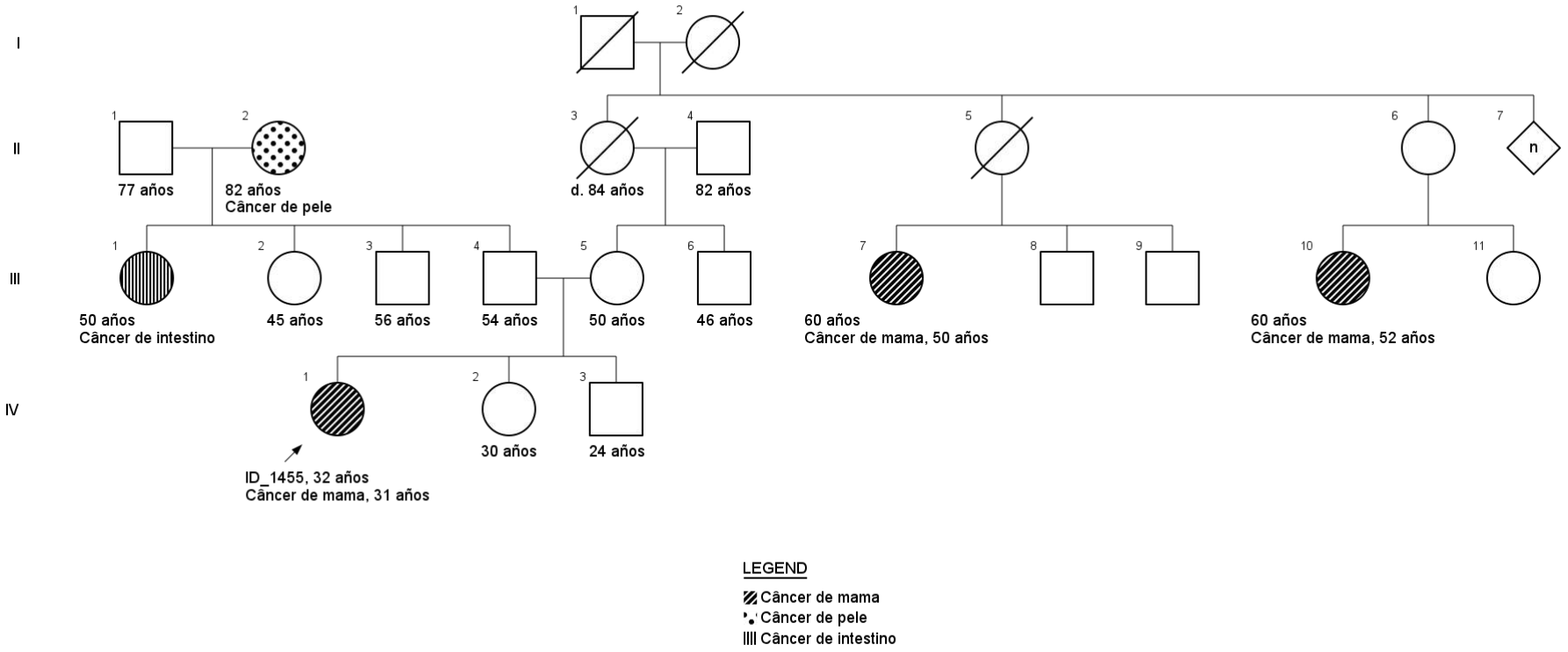
ANEXO LIII – Heredograma da paciente ID 1326.



ANEXO LIV – Heredograma da paciente ID 1327.



ANEXO LV – Heredograma da paciente ID 1455.



ANEXO LVI – Heredograma da paciente ID 1482.